

**WHAT ASPECTS OF REALISM AND
FAITHFULNESS ARE RELEVANT TO
SUPPORTING NON-VERBAL COMMUNICATION
THROUGH 3D MEDIUMS**

Simon CAMPION

School of Computing, Science and Engineering
The University of Salford, Salford, UK

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor
of Philosophy, June 2016

Table of Contents

LIST OF TABLES	VIII
LIST OF FIGURES	IX
LIST OF GRAPHS	XII
ACKNOWLEDGEMENTS	XIII
DECLARATION	XIV
ABBREVIATIONS	XV
ABSTRACT	XVII
1 CHAPTER 1 - METHOD	1
1.1 Introduction	1
1.2 Challenge	3
1.3 Scope	3
1.4 Approach	4
1.5 Research problem	4
1.6 Research Questions	5
1.7 Experiments	5

1.8 Contributions	6
1.8.1 Primary Contributions	7
1.8.2 Methodological Contributions	8
1.9 Impact	9
1.10 Publications	11
1.11 Overview of thesis	11
1.11.1 Research Question - How does the visual quality of CGI and video impact on identification of gross non-verbal behaviour in Virtual Chimpanzees?	11
1.11.2 Research Question - How do the visual qualities of video based 3D reconstruction and 2D video impact upon trust during lie detection and upon cognitive effort while doing so?	12
1.11.3 Research Question - How do the visual qualities of video based 3D reconstruction and 2D video affect the recognition of human facial expressions?	14
2 CHAPTER 2 - LITERATURE REVIEW	15
2.1 Methodology	15
2.1.1 Types of searches	16
2.1.2 Aim	16
2.1.3 Scope	16
2.2 Virtual Reality (VR)	17
2.2.1 Overview	17
2.2.2 Hardware and display technologies	21
2.2.3 Other relevant types of VR technology	25

2.2.4	Content Creation	27
2.2.5	Summary	28
2.3	Serious Games and VR applications	29
2.3.1	Overview	29
2.3.2	Examples	31
2.3.3	Serious games and virtual characters	32
2.3.4	Summary	33
2.4	Chimpanzee behaviour and virtual animals	34
2.4.1	Introduction to Chimpanzee behaviour	34
2.4.1.2	Chimpanzee hierarchy	36
2.4.1.3	Mimicking Behaviour	38
2.4.2	Zoo visitor education	39
2.4.3	Virtual animals	39
2.4.3.1	Virtual Pets	40
2.4.3.2	Virtual Primates	41
2.4.4	Summary	43
2.5	Virtual humans and human behaviour	44
2.5.1	Types of virtual character	47
2.5.2	Human behaviour	50
2.5.3	Challenges of representing non-verbal behaviour using a virtual character	55
2.5.4	Effect of virtual characters on users	64
2.5.5	Detection of truths and the fidelity of facial expression	67
2.5.6	Summary of the discussions in section 2.5	73
2.6	Summary	77

3	CHAPTER 3 - EXPERIMENT 1: GROSS BEHAVIOUR IDENTIFICATION – A COMPARISON OF VIDEO AND CGI FOR THE IDENTIFICATION OF CHIMPANZEE BEHAVIOUR.	82
3.1	Introduction	82
3.2	Method	83
3.2.1	Thesis and Hypothesis	83
3.2.2	Scope	84
3.2.3	Variables	84
3.2.4	Procedure	85
3.2.5	Subjects	86
3.2.6	Ethics	87
3.2.7	Apparatus	87
3.2.8	Questionnaire design	93
3.3	Results	94
3.3.1	Behavioural identification performance	94
3.3.2	Visual fidelity	95
3.3.3	Enjoyment	97
3.4	Discussion	98
3.5	Summary	101
3.5.1	Advantages of CGI visuals for portraying animal behaviour	101
3.5.2	Disadvantages of CGI visuals for portraying animal behaviour	102
3.5.3	Contributions of this experiment	102

4	CHAPTER 4 - EXPERIMENT 2: LIE DETECTION – A COMPARISON OF VIDEO AND 3D VIDEO FOR THE DETECTION OF DECEPTION	104
4.1	Introduction	104
4.2	Research Question	106
4.3	Hypotheses	107
4.4	Method	107
4.4.1	Scope	108
4.4.2	Session capture	109
4.4.3	Detection of Deception	114
4.4.4	Variables	116
4.4.5	Ethics	118
4.4.6	Participants	119
4.4.7	Analysis method	119
4.5	Results	119
4.6	Discussion	122
4.7	Summary	125
5	CHAPTER 5 - EXPERIMENT 3: FACIAL ACTION CODING – A COMPARISON OF VIDEO FRAMES AND 3D VIDEO FRAMES IN TERMS OF THE QUALITY OF FACIAL FEATURES AND EXPRESSIONS.	127
5.1	Introduction	127
5.2	Method	127

5.2.1	Identification of facial expressions	127
5.2.2	Variables	129
5.2.3	Procedure	129
5.2.4	Subjects	130
5.2.5	Participant environment	130
5.2.6	Analysis	130
5.3	Results	130
5.4	Discussion	131
5.5	Summary	133
6	CHAPTER 6 – DISCUSSION	134
6.1	Overview	134
6.2	Contributions	136
6.2.1	Contributions from the Literature Review	136
6.2.2	Contributions from Experiment 1: Gross behaviour identification – a comparison of video and CGI for the identification of Chimpanzee behaviour.	138
6.2.3	Contributions from Experiment 2: Lie detection – a comparison of video and 3D video for the detection of deception.	140
6.2.4	Contributions from Experiment 3: Facial Action Coding – a comparison of video and 3D video for quality of facial features and expressions.	143
6.3	Critical evaluation	146
6.3.1	Literature review	146

6.3.2	Experiment 1: Gross behaviour identification – a comparison of video and CGI for the identification of Chimpanzee behaviour.	146
6.3.3	Experiment 2: Lie detection – a comparison of video and 3D video for the detection of deception.	147
6.3.4	Experiment 3: Facial Action Coding – a comparison of video and 3D video regarding the quality of facial features and expressions.	148
7	CHAPTER 7 - CONCLUSIONS AND FUTURE WORK	149
	LIST OF REFERENCES	152
	APPENDICES	174
	Appendix 1: Information sheet and consent form for experiment 1	175
	Appendix 2: Chimpanzee Behaviours	177
	Appendix 2: Questionnaire for Experiment 1	178
	Appendix 3: Ethical Approval for Experiment 1	182
	Appendix 4: Recruitment email for Experiment 2	183
	Appendix 5: Information sheet and consent form for experiment 2 confederate	184
	Appendix 6: Information sheet and consent form for experiment 2 participant	186
	Appendix 7: Ethical Approval for Experiment 2	188

List of Tables

<i>Table 1. An overview of VR display technologies</i>	<i>22</i>
<i>Table 2. Table suggesting that full 3d reconstruction requires space and. generally more expensive hardware</i>	<i>59</i>
<i>Table 3. A comparison of virtual character attributes</i>	<i>61</i>
<i>Table 4. Independent Variables.....</i>	<i>85</i>
<i>Table 5. Independent variables</i>	<i>117</i>
<i>Table 6. Dependent Variables</i>	<i>118</i>
<i>Table 7. A table showing the correct identification of truth and deceit. The totals are out of a possible 126 statements.</i>	<i>120</i>

List of figures

<i>Fig. 1. Identifies the intersection of visual and behavioural realism and indicates the scope of the research.</i>	4
<i>Fig. 2. Venn diagram showing indication of spread of reading and the number of citations in each area.</i>	15
<i>Fig. 3. Inside the CAVE</i>	24
<i>Fig. 4. A Panoramic Display</i>	24
<i>Fig. 5. VICON Motion Capture Suite</i>	27
<i>Fig. 6. Left image illustrates Virtual human by Activision Blizzard 2012 (http://www.iryoku.com/blog), the right image is from Valves Half Life 2 from 2004 (http://half-life.wikia.com).</i>	47
<i>Fig. 7. Interpersonal distance between authored avatar and a virtuality avatar (Fairchild et al., 2015)</i>	51
<i>Fig. 8. Facial expressions and eye gaze portrayed by an abstract 3D avatar.</i>	52
<i>Fig. 9. Universal facial expressions of emotion and eye gaze. This figure shows a 3D reconstruction of the seven universal emotions and the quality of eye gaze in three directions.(Fairchild et al., 2015)</i>	53
<i>Fig. 10. Virtuality avatar waving. (Fairchild et al., 2015).</i>	55
<i>Fig. 11. This figure demonstrates the problem of occlusion (for the viewer) when using various display devices. Left: no stereo (full facial expressions); centre: stereo glasses (eye gaze and some facial features obscured) and right: HMD (most of the face obscured, identification of facial expressions not possible).(Fairchild et al., 2015)</i>	57
<i>Fig. 12 Depiction as to how the 3D reconstructed user is viewed from different viewpoints as one user moves around him. This free viewpoint navigation offers the user interactive</i>	

<i>functionality over and above that offered by traditional video conferencing, allowing users to view NVB from any angle (Fairchild et al., 2015)</i>	<i>58</i>
<i>Fig. 13. Virtual Character.....</i>	<i>88</i>
<i>Fig. 14. Virtual Chimpanzee</i>	<i>89</i>
<i>Fig. 15. Virtual Chimpanzee creation process from the motion capture (left) then via the untextured to the textured geometry.....</i>	<i>90</i>
<i>Fig. 16. Virtual Enclosure</i>	<i>91</i>
<i>Fig. 17. An image showing the virtual Chimpanzee enclosure</i>	<i>91</i>
<i>Fig. 18. Approach behaviour.....</i>	<i>92</i>
<i>Fig. 19. Eating behaviour.....</i>	<i>92</i>
<i>Fig. 20. Quadrapedal Walk.....</i>	<i>93</i>
<i>Fig. 21. The sender environment. The sender stands in the centre of the capture suite. A Macbook Pro connected to a webcam captures their image and transmits it through Skype. Simultaneously, multiple cameras (see Fig. 22) are recording and streaming images of the sender for 3D reconstruction.....</i>	<i>110</i>
<i>Fig. 22. Camera placement with respect to the sender position and display walls. Silhouette cones joins each camera to the silhouette edges of the sender as seen by that camera. Copyright IEEE (Roberts et al., 2013b).</i>	<i>110</i>
<i>Fig. 23. The environment shared by the questioner and the session display and capture. The 3D reconstruction was on the top screen and the video conferencing on the right. ..</i>	<i>111</i>
<i>Fig. 24. Participant environment. The participant is sat facing a desktop machine that is running the sender movies in E Prime. The participant is wearing an fNIRs sensor with a scarf wrapped around it to ensure light did not affect the data captured.</i>	<i>116</i>
<i>Fig. 25. Still frames showing video and virtuality avatar stimulus images.</i>	<i>128</i>

Fig. 26. Screen shot of virtuality avatar illustrating: 1. Slicing. 2. Lack of Concavities. 3.

The droop effect.....132

List of graphs

<i>Graph 1. The percentage accuracy of the correct identification of truth and lies. Reality values taken from Bond and DePaulo's meta-analysis (CF Jr Bond, 2006) and all other values taken from the study conducted by Steptoe et al. (Steptoe et al., 2010).</i>	69
<i>Graph 2. Showing behaviour identification.</i>	95
<i>Graph 3. Whether visual fidelity of CGI is high enough for behaviour identification.</i>	96
<i>Graph 4. Whether the video was more enjoyable than the CGI video.</i>	97
<i>Graph 5. Task performance - correct identifications of "truthfulness"</i>	120
<i>Graph 6. Comparing change in cognitive load across the two mediums - mean change in the oxygenation of blood in the prefrontal cortex for two groups</i>	121
<i>Graph 7. Mean change in oxygenation when viewing video representation against mean change in oxygenation when viewing virtuality avatar representation during the detection of truths</i>	122
<i>Graph 8. Accuracy of determining truthfulness across the media</i>	124

Acknowledgements

I would like to thank my supervisor Professor David Roberts for guiding me through this process and teaching me how to be a researcher, my manager Paul Welshman for allowing me the opportunity, and all of my work colleagues at VETS and THINKlab (with particular mention to Mikhail Polshaw, Michal Ceicura and Arturo Garcia) for support throughout. I would also like to thank my fellow PhD students Allen Fairchild, Tobias Duckworth and Carl Moore for allowing me to utilize their 3D reconstruction systems, John O'Hare for help with the OCTAVE system and from Psychology DR Peter Eachus, Robert Bendall and Aleksandra Landowska. For help and guidance with animal behavior DR Simon Hutchinson, DR Paul Rees, DR Sean O'Hara and Rebecca Dissan and Hannah Jarvis for advice and support on motion capture.

I would like to thank my wife Clare for encouraging and supporting me throughout, and putting up with the number of evenings and weekends I have spent working on this. My Parents, without them I couldn't have done any of it, and the rest of my family Rob, Sarah, Don, Abbie, Jodie, Sam, Jacob and my mother and father in law Mary and Michael all for keeping me smiling throughout. Plus anyone else who has helped me along the way, thank you.

Declaration

I declare that the work contained in this thesis is my own, and has not been submitted here or at any other higher education institution towards the award of a higher degree. Some of the work reported has been previously published.

Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AI	Artificial Intelligence
AR	Augmented Reality
AU	Action Units
CAVE	Cave Automatic Virtual Environment
CGI	Computer Graphics Imagery
FACS	Facial Action Coding System
fNIR	Functional near-infrared spectroscopy
FOV	Field of view
HCI	Human Computer Interaction
HMD	Head mounted display
ICVE	Immersive collaborative virtual environment

MOCAP	Motion Capture
MR	Mixed Reality
NPC	Non-player character
NVB	Non-verbal behaviour
NVC	Non-verbal communication
OCTAVE	Reconfigurable octagonal projection system
VE	Virtual Environments
VR	Virtual Reality
2D	Two Dimensional
3D	Three Dimensional

Abstract

This thesis investigates what aspects of realism and faithfulness are relevant to supporting non-verbal communication through visual mediums. The mediums examined are 2D video, 3D computer graphics and video based 3D reconstruction. The latter is 3D CGI derived from multiple streams of 2D video. People's ability to identify behaviour of primates through gross non-verbal communication is compared across 2D video and 3D CGI. Findings suggest 3D CGI performs equally well to 2D video for the identification of gross non-verbal behaviour, however user feedback points to a lack of understanding of intent. Secondly, ability to detect truthfulness in humans across 2D video and video based 3D reconstruction mediums is examined. Effort of doing this is measured by studying changes in level of oxygenation to the prefrontal cortex. Discussion links to the literature to propose that tendency to over trust is inversely proportional to the range of non-verbal resources communicated through a medium. It is suggested that perhaps this is because "tells" are hidden. The third study identifies that video based 3D reconstruction can successfully illustrate subtle facial muscle movements on a par with 2D video, but does identify issues with the display of lower facial detail, due to a reconstruction error called droop. It is hoped that the combination of these strands of research will help users, and application developers, make more informed decisions when selecting which type of virtual character to implement for a particular application therefore contributing to the fields of virtual characters and virtual environments/serious gaming, by giving readers a greater understanding of virtual characters ability to convey non-verbal behaviour.

1 Chapter 1 - Method

1.1 Introduction

This PhD initially started in order to examine the use of virtual reality to represent animal behaviour, in particular, Chimpanzee behaviour. However, part way through the study it was realised that capturing more subtle behaviour, such as facial expressions, where markers need to be applied to the face, or a subject needs to be put within a capture volume are not viable for animal welfare and safety reasons. As a result of this, it was decided to switch focus to human behaviour for the latter sections of this research which involved more subtle non-verbal behaviour. Due to similarities in Chimpanzee and human facial musculature, it is suggested that if the quality of virtual character facial expression is sufficient for human representation, then it should illustrate the concept for Chimpanzees. An additional benefit of expanding the study to cover human behaviour was that it makes the PhD more interesting to a wider audience, to those in academic fields and to developers. The resultant PhD is multidisciplinary and draws upon the areas of virtual environments (VE), serious games, animal behaviour and human behaviour.

In Virtual Reality (VR), the quality of virtual characters can be measured using scales such as realism and believability. Realism can be broken down into several sections: visual, behavioural, acoustic, smell and haptic. This study focused on visual and behavioural realism and its impact on the identification and interpretation of non-verbal behaviour, with the aim of investigating what this researcher will term 'relevant realism'. If it is possible to identify the optimal level of visual and behavioural realism for the identification and

interpretation of behaviour, characters will be believable, fit for purpose, and more cost effective to produce.

Virtual reality can induce realistic reactions in people through providing simulation that is believable, without needing it to be realistic. However, while believability might be sufficient in a game, would a lack of visual and behavioural realism impact on meaning in real conversations and more serious applications, and thus on trust? Compelling experimental environments and CGI videos depicting behaviours have been created to test hypotheses. The first stage of this research examined gross full body movements in Chimpanzees, the second the more subtle facial expressions in humans.

It is known that VR is capable of providing believable simulations and that users can be represented as avatars. However, in many past studies into immersive collaborative virtual environments (ICVE), these avatars have traditionally been constructed of separate body parts and have offered a limited range of behaviour, often limited to head and hand tracking which could be hindered by tracking issues. Game engines have improved things somewhat by allowing researchers to utilise more realistic full body avatars. A single mesh represents the character and it is rigged with a skeleton which is used to deform the mesh and allow more realistic animations. However, it is challenging to simultaneously capture full body and more subtle facial movements together. The evolution of video-based reconstruction offers solutions to some of these problems allowing the capture of both full body and facial animations. This research explores these mediums and their ability to convey non-verbal behaviour. It is important to be able to identify a behaviour, but it is equally necessary to identify the intent behind it. Depending on the application required? varying levels of realism and believability are required to make an agent or avatar fit for purpose.

1.2 Challenge

The main challenge of this research is to explore where visual qualities impact on perceived behavioural realism and interpretation of meaning.

1.3 Scope

This PhD began by surveying the literature on social agents and virtual characters and examining studies of social interactions between animals, between animals and humans, and between humans. The aim of this thesis was to investigate what aspects of realism and faithfulness are relevant in supporting non-verbal communication through visual mediums. In order to explore this it was necessary to break behaviour down into smaller portions, allowing the study of the identification of gross full body movements, subtle body movements such as facial expressions, and the intent or meaning derived from their identification. This thesis examines both virtual animals and virtual humans and thus contributes information that will be of use to both fields.

For the first stage of the study a prototype was developed that used motion tracking to capture full body animal behaviour (Chimpanzee) for the experimental environment. Virtual characters were created and brought to life through interactive behavioural animation, then placed in a virtual environment. Following this, the focus switched to more subtle human behaviour with facial movements being captured using video-based reconstruction, allowing studies of subtle avatar representations of facial behaviour. Evaluation included both subjective and objective measurements of participants identifying behaviour. The identification and interpretation of behaviour afforded by the simulation was compared to

that of a 2D video medium in an effort to identify levels of relevant realism. Fig. 1 illustrates the intersection of visual and behavioural realism, indicating the scope of this research.

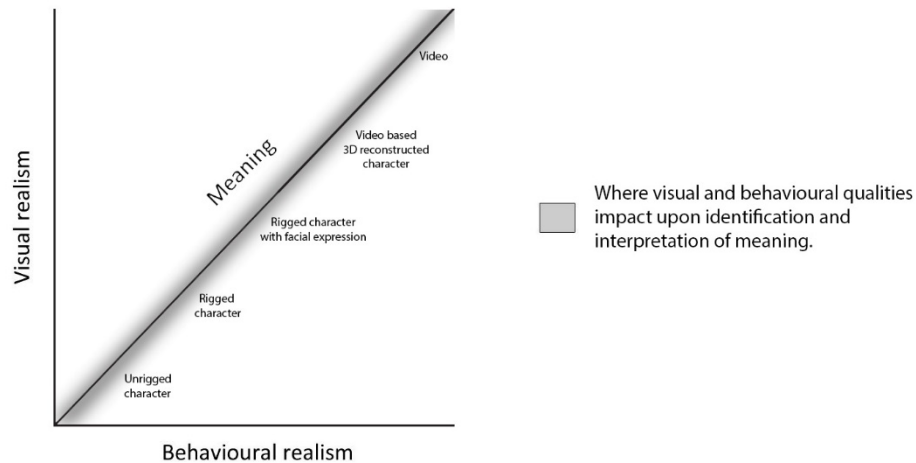


Fig. 1. Identifies the intersection of visual and behavioural realism and indicates the scope of the research.

1.4 Approach

This research began with a literature survey identifying gaps in knowledge. A series of experiments were then designed with the aim of adding knowledge in these areas. Finally, the results are discussed, conclusions made and potential future work outlined.

1.5 Research problem

How is the identification and interpretation of behaviour impacted by differences in visual and behavioural realism across 2D video and 3D media?

1.6 Research Questions

The following research questions have been designed to build upon one another in order to explore the identification and interpretation of behaviour through virtual characters.

1. How does the visual quality of CGI and video impact on the identification of gross non-verbal behaviour in Virtual Chimpanzees?
2. How do the visual qualities of video-based 3D reconstruction and 2D video impact upon trust during lie detection and upon cognitive effort while doing so?
3. How do the visual qualities of video-based 3D reconstruction and 2D video affect the recognition of human facial expressions?

1.7 Experiments

- 1. How does the visual quality of CGI and video impact on identification of gross non-verbal behaviour in Virtual Chimpanzees?**

When exploring the use of virtual environments, researchers often become interested in interaction and ignore the other components of the environment, thus little research has been conducted simply comparing the visual variables. The aim of this part of the study was to remove all interaction and look at how virtual environments, as a purely visual medium, compare with video for the identification of gross full body Chimpanzee behaviour.

2. How do the visual qualities of video based 3D reconstruction and 2D video impact upon trust during lie detection and upon cognitive effort while doing so?

Is trust impacted by the visual differences of 2D video and video based 3D reconstructed avatars, and what is the effect on cognitive effort? This has implications for many applications that use avatars, such as educational virtual environments and social media.

3. How do the visual qualities of video based 3D reconstruction and 2D video affect the recognition of human facial expressions?

Building on the data set from experiment two, this is a study of how a video based 3D reconstructed avatar representation impacts on the recognition of facial expressions. Can a 3D reconstructed video faithfully convey facial information as well as a video. Is it close enough for it to be a useful training tool? A high percentage of behaviour is non-verbal, so if virtual characters do not portray these behaviours adequately are they yet fit for purpose?

1.8 Contributions

There has been much research into the realism of virtual characters. In this thesis the focus has been on the identification and interpretation of non-verbal behaviour through video and 3D mediums examining both gross full body behaviour and more subtle facial movements. Video based 3D reconstruction gives a new and exciting approach to character creation allowing the fluid capture of both full body and more subtle non-verbal behaviour through simultaneously capturing the geometry and textural information of a form. Authored CGI and video based 3D reconstructed avatars have been studied to better understand their

qualities, strengths and weaknesses. The overall contribution to science from this research is adding knowledge to the area of virtual characters and, in particular, their ability to portray non-verbal behaviour, examining how such representation impacts on behaviour identification and on our understanding of the meaning behind them.

1.8.1 Primary Contributions

1. The first contribution was the undertaking of a new literature survey identifying the following gaps in knowledge:
 - i. Identifying a possible trend/opportunity across the literature that suggests that zoos need a better way to inform visitors about animal behaviour thus contributing to serious gaming and conservation education.
 - ii. Identifying a possible gap and opportunity in the literature that suggests that zoologists would benefit from better ways of learning about animal behaviour.
 - iii. Identifying a possible trend across the literature that suggests that reproducing facial expressions reduces the tendency of observers to over trust. However, the researcher could only identify a limited number of studies in this area, so more work would be needed to confirm this.
2. The second contribution was demonstrating that CGI with low level non-verbal behaviour was sufficient for the identification of gross NVB on a par with video.

3. The third contribution was identifying a possible issue with the identification of behavioural intent due to a lack of subtle NVB.
4. The fourth contribution was demonstrating, through an experiment, that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of deceit.
5. The fifth contribution was showing, through an experiment, that participants work equally hard when detecting deceit through a 3D reconstructed medium and video conferencing.
6. The sixth contribution was linking the above two contributions (contributions four and five) to provide additional evidence to support the above trend.
7. The seventh contribution of this work was to demonstrate that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of facial muscle movements.

1.8.2 Methodological Contributions

1. Methodological contribution – Motion tracking of a zoo keeper mimicking primate behaviour.
2. Locking and matching viewpoints to allow an exact contrast of visual quality without the distraction of interaction.

3. Experimental setup for simultaneously capturing 2D video and video based 3D reconstructed video for contrast purposes.
4. The use of fNIRS to compare cognitive effort during the detection of truths in 2D video and video based 3D reconstruction for contrast purposes.
5. The use of facial action coding to measure the quality of video based 3D reconstruction when compared to 2D video.

1.9 Impact

This thesis builds upon research at The University of Salford conducted by the VR and Telepresence group led by Professor David Roberts. Utilising the video based 3D reconstruction system developed and refined by a series of researchers, it is hoped that experiments reported in this thesis will feed back into the future development of this system, improving quality and usability, while demonstrating its potential for application in other academic disciplines and commercial activities.

The first experiment of this research illustrated that authored abstract 3D avatars are capable of displaying gross non-verbal behaviour on a par with video but suggests that, while this is possible, it is difficult for users to gauge intent due to a lack of subtle behaviours. In the second experiment, a comparison of video and video based 3D reconstruction was conducted during the identification of truths. There was no significant difference between identification across the two media. In addition, it could be seen that the visual differences between the media do not alter how hard the viewer works when identifying truths. Comparison with the literature suggests that selecting an avatar with less behavioural facial detail could possibly effect truth bias, yet, due to the relatively small number of studies, more work would be

needed to prove this. However, if it is true, it has implications for serious applications and social media. The third experiment of the research showed that video based 3D reconstruction could display facial muscle movements on a par with video with the exception of lower facial action units.

These findings suggest that, for applications where avatar mediated social interaction is required, developers may need to devote more thought to matching behavioural realism to application. If a user selects an avatar that is unable to display subtle non-verbal behaviour, then other users may find it difficult to identify the intent of their actions. Viewers of avatars, or indeed agents, may assume they are getting all the relevant communication from the character when, in essence, they may be getting a poker face in all situations. This research is also of interest to the VR, gaming and serious gaming community where developers may be deciding how much time, effort and detail to put into characters. This study helps them to identify that time spent selecting the correct level of relevant realism will ensure that their characters depict all the information needed to provide believability for a particular task, therefore assisting in meeting their learning outcomes whilst not wasting time on unnecessary behavioural detail.

It is hoped that this research will aid developers and users of virtual characters by illustrating that it is not necessarily enough to simply be able to identify behaviour. The research suggests it is equally important to be able to understand intent. In addition, the research highlights some strengths and weaknesses of abstract and video based 3D reconstructed avatars. It is again hoped that this may help developers and users choose the correct medium.

The thesis could be of interest to VR researchers, psychologists, criminologists or behavioural therapists interested in using or studying various 3D media's ability to convey non-verbal behaviour. Potential application areas include a procedural trainer for people working with animals, an educational tool used for conservation education, a teaching aid for social literacy and a procedural trainer for law enforcement and security personnel.

1.10 Publications

1. A comparison of Video and CGI for the identification of Chimpanzee Behaviour.
(Presented at Games and Learning Alliance (GALA) 2015, to be published in Springer Lecture Notes in Computer Science)
2. Removing the mask - do people over trust avatars reconstructed from video? (Under Review, first Author.)
3. A Mixed Reality Telepresence System for Collaborative Space Operation (Under Review, Second Author.)
4. withyou—An Experimental End-to-End Telepresence System Using Video-Based Reconstruction (IEEE Journal of Selected Topics in Signal Processing, Third Author.)

1.11 Overview of thesis

- 1.11.1 **Research Question** - How does the visual quality of CGI and video impact on identification of gross non-verbal behaviour in Virtual Chimpanzees?

Contribution

1. The first primary contribution was in identifying a possible trend across the literature that suggests that zoos need a better way of informing visitors about animal behaviour, contributing to conservation education.
2. The second was identifying a possible gap in the literature that suggests that zoologists would benefit from a better way of learning about animal behaviour.
3. The third contribution was demonstrating that CGI with low level NVB is sufficient for the identification of gross NVB on a par with video.
4. The fourth contribution was identifying a possible issue with the identification of behavioural intent due to a lack of subtle NVB.

Methodological contribution – Motion tracking a zoo keeper to mimic primate behaviour.

Publication

A comparison of Video and CGI for the identification of Chimpanzee Behaviour.

(Presented at GALA 2015, to be published in Springer Lecture Notes in Computer Science)

Chapter

Chapter 3 - Experiment 1: Gross behaviour identification – A comparison of Video and CGI for the identification of Chimpanzee behaviour. Page 82.

1.11.2 **Research Question** - How do the visual qualities of video based 3D reconstruction and 2D video impact upon trust during lie detection and upon cognitive effort while doing so?

Contribution

- 1 The first contribution was demonstrating, through an experiment, that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in overall portrayal of deceit.
- 2 The second contribution was showing, through an experiment, that participants work equally hard when detecting deceit through a 3D reconstructed medium and video conferencing.
- 3 The third contribution was linking the above two contributions to provide additional evidence to support the above trend.
- 4 The first methodological contribution was locking and matching viewpoint to allow an exact contrast of visual quality without the distraction of interaction.
- 5 The second methodological contribution was the experimental setup for simultaneously capturing 2D video and video based 3D reconstructed video for contrast purposes.
- 6 The third methodological contribution was the use of fNIR to compare cognitive effort during the detection of truths in 2D video and video based 3D reconstruction for contrast purposes.

Publication

Removing the mask - do people over trust avatars reconstructed from video? (Under Review, first Author.)

Chapter

Chapter 4 - Experiment 2: Lie detection – A comparison of video and 3D video for detection of deception. Page 104.

1.11.3 **Research Question** - How do the visual qualities of video based 3D reconstruction and 2D video affect the recognition of human facial expressions?

Contribution

1. The primary contribution of this work was demonstrating that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of facial muscle movements.
2. The methodological contribution was the use of facial action coding to measure the quality of video based 3D reconstruction when compared to 2D video.

Publication

Removing the mask - do people over trust avatars reconstructed from video? (Under Review, first Author.)

withyou—An Experimental End-to-End Telepresence System Using Video-Based Reconstruction (IEEE Journal of Selected Topics in Signal Processing, Third Author.)

Chapter

Chapter 5 - Experiment 3: Facial Action Coding – A comparison of video frames and 3d video frames for quality of facial features and expressions. Page 127.

2 Chapter 2 - Literature review

2.1 Methodology

This study is multidisciplinary and draws upon knowledge from several different areas including virtual reality, serious games, animal behaviour and human behaviour. The Venn diagram (Fig. 2) below illustrates the split between these subjects and where they overlap.

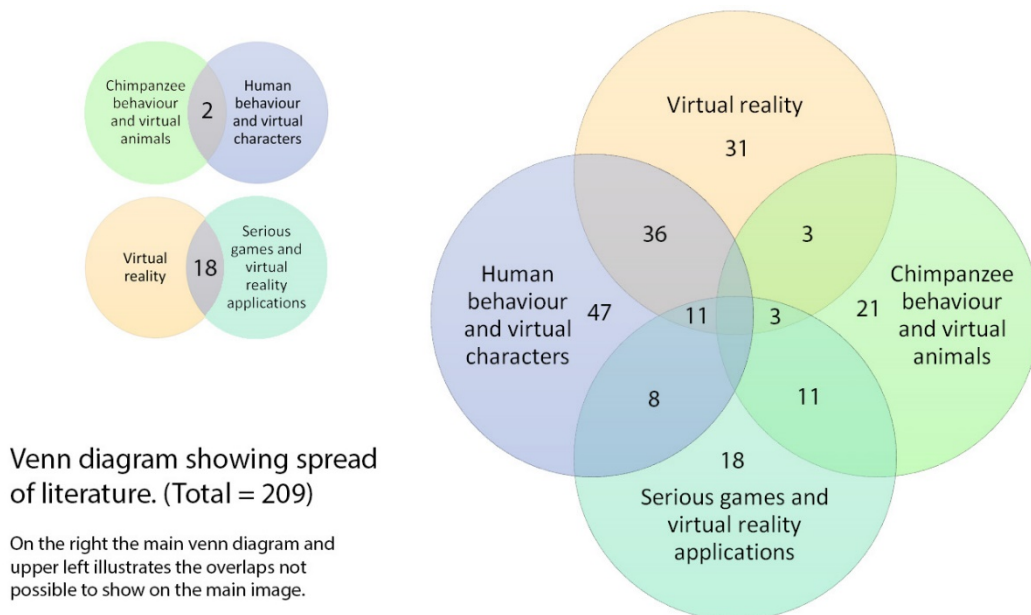


Fig. 2. Venn diagram showing indication of spread of reading and the number of citations in each area.

2.1.1 Types of searches

The literature review was constructed from many sources including books, journals, conference papers, expert interviews and video footage. These sources were located using library resources, online searches and interviews.

2.1.2 Aim

The aim of the review was to survey the literature on social agents and virtual humans, look at the studies on social interactions between animals, animals and humans, and between humans, examining virtual characters' current ability to portray non-verbal behaviour, and to identify gaps in knowledge where further research is required.

2.1.3 Scope

In order to develop an overall understanding of virtual characters, this review will be split into four areas:

- i. The first area will introduce virtual reality (VR) and its background. VR, or a virtual environment (VE), is the habitat of virtual characters. It is important to understand its composition in order to create virtual characters that can fit and function within it. Within this section various VR hardware and display systems will be discussed, with a view to understanding their ability to display VEs. Including their ability to track and display user behaviour and problems that can arise when doing this.

- ii. The second area of the review will cover serious games (educational virtual environments), an area that is growing in popularity. It is here that the full repertoire of a virtual character's functionality may be of most use. Developers and researchers often employ virtual characters in roles such as teaching, guidance or storytelling.
- iii. The third area of this review will concentrate on animal behaviour with an emphasis on NVB, focusing on Chimpanzees and, finally, on exploring virtual animals.
- iv. The final area covered within the review will be human behaviour and virtual characters, exploring types of non-verbal behaviour (NVB) that could be incorporated into virtual character design, examining how virtual characters are made, what they can do and where improvements could be made.

This thesis was initially going to focus primarily on primate behaviour, in particular on Chimpanzees. However as it developed, it became apparent that expanding the focus to include human behaviour allowed the study of not only more subtle behaviours, but also the examination of new, emerging virtual character technologies and techniques, such as video based 3D reconstruction.

2.2 Virtual Reality (VR)

2.2.1 Overview

The first recorded use of the term virtual reality was by Jaron Lanier (Conn et al., 1989). There are now many definitions of VR, one such being the Oxford Dictionary definition which is “The computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special

electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors” (oxforddictionaries.com). Steuer suggests that people often talk about VR in a similar manner to a television or telephone, concentrating on the hardware such as head mounted displays (HMD) and tracking systems rather than on user experience (Steuer, 1992). This is a valid point. In computer gaming it is only since the emergence of the commercial HMDs that people have started talking about VR games. However, the only thing that has changed is the display technology, as the majority of the development remains the same. The term VR has been defined many ways by different academics, the standpoint of this thesis is that VR is a computer generated interactive environment that simulates real places and situations, allowing the user to navigate and interact with the environment in real time.

When people talk about VR they often talk about the terms Presence and Immersion. These terms have been subject to much discussion with academics interchanging both. For the purpose of this study, the thesis adopts Slater’s definitions: “The more that a system delivers displays (in all sensory modalities) and tracking that preserves fidelity in relation to their equivalent real-world sensory modalities, the more that it is 'immersive'.” (Slater, 2003). Slater describes presence as follows: “Presence arises from an appropriate conjunction of the human perceptual and motor system and immersion. Presence is a response. Separate from presence are aspects of an experience such as involvement, interest and emotion. These are to do with the content of the experience. Presence is the form.” (Slater, 2003). To summarise these definitions for this thesis: Immersion in virtual reality is the level to which stimuli from reality is exchanged for the virtual, and can create a feeling of ‘being there’, referred to as presence. With regards social interactions, social presence, or the feeling of sharing an environment with another, is defined as “the moment-by-moment awareness of

the co-presence of another sentient being accompanied by a sense of engagement with the other” (Biocca et al., 2001, Bente et al., 2008).

In its early years, VR was perhaps a victim of its own potential. The media, and films such as *The Lawnmower Man* (Leonard, 1992), gave people false expectations of what was attainable at that time. This combined with the excessive costs of VR hardware and the expertise needed to create even simple environments, slowed the area’s progress and commercialisation. In recent years, however, things have changed. The entertainment industry has embraced 3D graphics in computer games, films and television. Vast budgets are spent on 3D motion pictures and large video game titles. In order to play the latest games, consumers need to keep up to date with hardware. To maximise uptake, this hardware was aimed at the mass markets and, therefore, had to be affordable. This driving down of hardware costs made it possible to create rich 3D environments on gaming PCs, provided that software and expertise are available. Simultaneously, software improvements have been constant with large game and independent studios releasing versions of their rendering engines for use by smaller developers. The combination of cheaper, powerful hardware and ‘off the shelf’ software, have been the catalyst for a boom in 3D graphics uptake.

It is not just computer graphics’ rendering and standard hardware that have seen improvements and lowering costs. It has also happened with interaction devices. In the past, the use of motion capture equipment would have been an expensive option. Now, Microsoft Kinect (Kinect) and other similar devices are making it available to the masses. In addition, previously, HMDs were large, heavy and expensive. They fell out of fashion and researchers moved towards larger scale immersive displays. Today, however, HMDs are being commercialised for VR and augmented reality (AR) use, with devices such as the Oculus

Rift (Oculus), Google Glass (Google, 2015) and Hololens (Microsoft). Many of these new devices are remakes of older VR technologies, now available to everyone at a reasonable cost.

For a time the term VR was not fashionable. Now, its concepts are starting to emerge under various names and brands. This is positive for the research community, because researchers now have reasonably priced hardware available to conduct studies. Once studies are complete, it will be easier to conduct technology transfer to industry without excessive hardware costs. These factors are important because it is this momentum which is making research into VR very important at this time. As VR starts to enter homes, workplaces and educational institutions, researchers and developers need to ensure that content is fit for purpose.

VR simulations are interactive and can afford high levels of presence. Of course, as suggested by Steuer (Steuer, 1992), a sense of presence can be gained from a variety of media, from books and films through to VR. Steuer uses interaction and vividness as variables to measure presence, suggesting that higher levels of interaction and vividness result in higher levels of presence. However, there can be exceptions. A movie, for example, would score highly in terms of vividness but badly on the interactive scale, and would, therefore, offer a lower sense of presence than that of an interactive VE for instance. In addition to the link between levels of immersion and presence, there are also links between realism and presence. Realism can be structured using our sensory channels, thus has several sub-topics: visual realism, acoustic realism, scent realism, realism of taste and haptic realism. Studies have shown that increasing the level of visual realism raises the level of presence afforded to the user (Khanna et al., 2006), further evidenced by changes in

participant's physiological data in more realistically rendered environments (Slater et al., 2009). Both studies compared the use of ray casting and ray traced rendering techniques. Participants using the more realistic ray traced environment exhibited higher levels of presence. However, it is possible to be fully immersed within VR hardware running low fidelity graphics and feel a sense of presence (Slater, 2003), raising interesting questions about how much detail is needed to achieve a presence response.

2.2.2 Hardware and display technologies

This section of the review examines what options are available for VR deployment. These range from high-end fully immersive installations to desktop PC solutions. It is important to understand what options are available, as the amount of immersion will impact on the levels of presence afforded to the user. Display technologies can be broken down into types. As previously explained, immersion in virtual reality is the level to which stimuli from reality are exchanged for the virtual. This is achieved using various hardware and software combinations. The more the user's senses can be engaged, the higher the level of presence should be. HMDs have typically been associated with VR, however, large scale projected display technologies are also popular. Examples include the CAVE (Cruz-Neira et al., 1992) and the OCTAVE (O'Hare, 2015), as they are considered to be spatially immersive display technologies (Lantz, 1996). They allow multiple people to walk into the display space, yet head tracking for real-time viewpoint update is limited to one user. Therefore, all users will see the world from the tracked user's visual perspective. Viewpoint update is achieved using tracking systems such as those from Vicon (Vicon, 2015). Typically, markers are placed on stereo glasses in a set pattern. This pattern is identified by the tracking system allowing translation and rotation to be recorded in real-time. This information is then used to update

the projected viewpoint of the virtual world to match the movement of the wearer. When tracking systems are used in conjunction with stereoscopic displays, such as CAVEs, it enables the tracked user to look or walk around a virtual object or character.

Visual display type	~ Indicative view angle (Horizontal (H) and Vertical (V))	Cost	Tracking (Degrees of freedom rotation and translation)	Stereoscopic available
HMD	360 H 360 V	Low	Head tracking and full body with optical tracking	Yes
CAVE/OCTAVE	270 H 200 V for 3 wall and floor CAVE. 360 H 200 V for OCTAVE.	High	Possible with full body optical tracking	
Panoramic screen	160 H 70 V	Medium		
Power wall	90 H 90 V	Medium to high		
Hemespheric display	180 H 180 V	Medium to high		
Projection	90 H 90 V	Low to medium		
Monitor	90 Horizontal 90 Vertical	Low	Possible upper body with optical tracking systems.	

Table. 1. An overview of VR display technologies

A brief description of typical VR display devices is shown below.

HMD – A head mounted display allows the user a visually, fully immersive experience. The 3D graphics are drawn on a small screen, or screens, positioned in front of the user's eyes consuming their field of view (FOV). When a user moves his/her head, it is motion tracked and their view of the environment is updated accordingly. This technology can only be used by one person at a time. However, it is possible to use these collaboratively allowing remote participants to have a shared experience. A potential problem for this technology is that current consumer HMDs such as the Oculus, VIVE and Morpheus (Oculus, HTC, 2015, Sony, 2015) cover most of the user's face, meaning capture and display of facial expressions is not possible at this time.

CAVE/OCTAVE – The CAVE (Cruz-Neira et al., 1992) and the OCTAVE (an eight sided immersive display) (O'Hare, 2015) are fully immersive environments, offering full stereo 3D graphics on four and eight screens respectively, allowing the user to stand in the middle of a computer generated VE. These systems are equipped with motion tracking technology to capture user movement for interaction. Such facilities do allow for more than one person to be present but only one person will be tracked at a time, and the rendered view of the VE will be updated according to that individual's visual perspective. The downsides of this technology are cost and space. These systems typically use immersive collaborative virtual environment (ICVE) avatars. However, newer systems can sometimes capture and display video based 3D reconstructed avatars.



Fig. 3. Inside the CAVE

Panoramic display – A panoramic display is a curved semi-immersive display that takes in more of a user’s FOV than a flat screen. The system pictured in Figure 4 uses 150 degrees of the user’s FOV when sat front and centre. It is well suited for demonstrations to small groups of individuals and can fit up to twenty people. The downside to this technology is, again, the cost as the curve in the screen generally means multiple projectors and edge blending hardware are required. Today television sets using similar curved designs are starting to be seen (Samsung, 2015).



Fig. 4. A Panoramic Display

Hemispheres - A hemisphere is semi-immersive half sphere screen into which the VE is projected. These can be large or small scale. When produced on a larger scale they can again be costly because of the edge blending hardware or because a more costly lens is needed for the projection (ImmersiveDisplay, 2015).

Power walls - A power wall can be constructed of an array of monitors or projection tiles that allow content to be shared across them. They produce similar results to a large flat projection, but provide a brighter image and a higher resolution than standard projectors. The downside of this technology is the visible joins between displays and, depending upon the individual units used, they can be costly (Christie, 2015).

Standard data projectors – Although not as high end as some of the other display solutions mentioned, a correctly set up high resolution projector can still provide a more than adequate visualisation solution. The advantage of a data projector is that the cost is relatively low.

Monitor – The use of standard display monitors to run Virtual Environments allows a larger group of people to have access to the use of VEs. In addition, a stereo monitor allows the user to gain depth cues from 3D graphics.

2.2.3 Other relevant types of VR technology

Mixed reality (MR) and Augmented reality (AR) - Augmented reality is when 3D graphics are layered onto a real scene. This can be achieved using Augmented Reality goggles (Microsoft). Many systems use pattern recognition software that identifies markers and overlays the 3D content onto their positions. AR has many exciting potential

applications within training. Information or processes can be layered onto a user's vision in real time as he/she completes tasks. Mixed Reality is described as a continuum by Milgram and Kishino (Milgram and Kishino, 1994), with differing levels of virtual and real information placing an application somewhere upon the continuum.

Telepresence/3D reconstruction systems - Telepresence is the creation of 3D avatars from real-time video streams (Roberts et al., 2015a). Systems can capture and remotely present users as avatars in another location. Multiple cameras are used, positioned at different angles around the subject to be captured. Common methods of 3D reconstruction include shape from silhouette (Roberts et al., 2013a) and depth-based approaches (Maimone and Fuchs, 2011, Kinect). The resulting remote 3D representation of a user can be viewed from any angle with a free viewpoint update. Telepresence has many potential real-time application areas such as video conferencing, meetings and teaching, while offline rendering could make the technology useful for performances or films.

Motion Capture/Tracking systems - Motion Capture (MOCAP) can be used in two ways. Firstly, to capture data which can then be processed offline for later use in animation. The second use is to capture data in real-time for user tracking and embodiment in systems such as CAVEs (Cruz-Neira et al., 1992). MOCAP can also be used for other purposes such as sports science and medical applications. "Motion Capture (MOCAP) is sampling and recording the motion of humans, animals and inanimate objects as 3D data. The data can be used to study motion or to give an illusion of life to 3D computer models." (Kitagawa and Windsor, 2008).



Fig. 5. VICON Motion Capture Suite

2.2.4 Content Creation

Typical content creation methods are outlined below.

Modelling – VR modelling techniques utilised by industry and academia are, by and large, now the same as those employed for the creation of computer games; that is, employing low polygon techniques to create geometry with optimal triangle counts enabling faster real-time rendering speeds. However, hardware specifications have improved consistently year on year so model specifications have increased. Examples of modelling packages are 3DS Max (Autodesk, 2015a), Maya (Autodesk, 2015b) and Blender (Blender, 2015). Other model creation techniques include laser scanning and photogrammetry.

Texturing – Texturing plays its part in optimising model performance. Detail can be added to models using normal or bump maps by making a low polygon model appear as a high

polygon model. Lighting information can be baked into textures in order to avoid calculations at run time, while specular and reflective qualities can also be saved into a texture map. Adobe Photoshop has been used for this study and is an industry standard for texture creation (Adobe, 2015).

Rigging – Rigging is where a skeleton or bones are applied to virtual character geometry. It is created and scaled to fit the geometry and then assigned to it. Once the geometry is linked, each vertex is assigned and weighted to a particular bone/bones so that the skin of the character deforms correctly when animation occurs.

Animation – Animations can be created using key frames, motion capture data or physics' engines, giving the animator the ability to simulate processes, characters and events.

Interaction - Game engines are becoming more widely accepted in Virtual Reality research. Just a few years ago, any visualisation or interactions would require a great deal of programming expertise but now off-the-shelf products such as Quest3D (ACT3D, 2015), Unity3D (Unity3D, 2015) and the Unreal Engine (EpicGames), to name just three, are making interactions much easier to achieve, without high level coding skills. However, many people still choose to programme when creating their resources, giving them full control over the scene graph and the rendering pipeline.

2.2.5 Summary

In summary now is an exciting time for VR. Hardware and software have become much more accessible thanks to dedicated researchers, developers and a lot of investment from the

games and entertainment industries. New hardware has provided researchers with opportunities to develop research on affordable systems while software improvements and, in particular, game engines have freed researchers from large amounts of programming, thus giving them more time to concentrate on creating richer experimental VEs and, more importantly, has allowed them to concentrate on research problems.

Reduction in hardware costs has made VR more mainstream. This year consumer HMDs will be released for the masses, and with them VR games and environments. Users within these environments will be represented by avatars with varying levels of NVB. If these environments are collaborative, it would be desirable to know how well virtual characters, both avatars and agents, portray NVB and whether users can identify intent. The level of NVB available to users is dependent on software and hardware; therefore, it was important for this research to review the technology currently available.

2.3 Serious Games and VR applications

2.3.1 Overview

Serious games are games that are created not for entertainment but for other applications (Susi et al., 2007). They apply computer gaming techniques and technology to learning, enabling a more interactive and engaging experience. Multimedia content can be embedded within these games such as video, audio, documents, mini games or simple web links, transforming the environment into an interactive hub of knowledge, wherein the player can explore subjects in a playful way that engages multiple senses rather than searching the web and getting lost in a mass of information (Njenga and Fourie, 2010). There is also a

crossover into mainstream gaming with the emergence of many 'edutainment' titles that cover subjects from maths and cookery to exercise (Shogakukan, 2008, Ubisoft, 2008, Nintendo, 2007). Serious games, such as those mentioned, could revolutionise the way in which subjects are taught, with students learning new skills through exploration and play (De Freitas et al., 2010). More 'hands on' subjects, such as joinery, could use animation to show processes. Research into serious games suggests that they will enable teaching materials to be updated, adding to traditional teaching methods (Pan et al., 2006b) or, in some cases, be an alternative to text books (Hew and Cheung, 2010).

Using serious games opens up many new opportunities for training, much of which would either not be possible or would be too expensive to act out in real life (Ieronutti and Chittaro, 2007, Falloon, 2010, Dalgarno and Lee, 2010). A good example of this would be emergency service training applications. It would be costly for the police to hire actors every time they want to conduct an interview training, but they could use a virtual reality application as discussed by Frank et al. (Frank et al., 2002). The fire service cannot set a building on fire every time they run a training session, or run large scale scenarios requiring many fire fighters and various appliances. They can, however, run a serious game that simulates all of this (XVR, 2015). With the use of database back ends to store scenario events and interactions, information from training scenarios can be stored and integrated for performance. The training can then be re-run any number of times or, by changing a few variables, the scenario can be executed in an entirely different way. Rizzo and Kim (Rizzo and Kim, 2005) state that, in some instances, VR is not required to exactly copy reality. In other words, reality has constraints, such as time and safety. With VR these constraints can be put aside and scenarios can be played and repeated in any place and time.

2.3.2 Examples

Below are some examples of serious games.

1. Active Worlds – The River City Project (For American Schools, Middle Grades 6 to 9)

In this serious game, students must time travel back to the 19th century. They can then use a 21st century skill set to solve 19th century issues. The students work in teams to decide what health problems exist by recording clues. They then identify the condition and produce experiments to test their diagnoses (Worlds, 2015) (Ketelhut and Schifter, 2011).

2. Providing Geographically Dispersed Students and Faculty a place to meet and learn together: The Open University in Second Life

The Open University offer Higher Education through Distance Learning. Therefore, students do not get as many opportunities to meet, exchange ideas or collaborate as those attending University full time. In an effort to facilitate these kinds of interactions, the Open University has created a Virtual Campus in Second Life. Classes are held in virtual classrooms, or on islands dedicated to topics, such as Virtual Hallucination Island. On this island, Psychology students can experience some mental illnesses or disorders in the same way that a patient would (Life, 2015).

3. The VIBE

The VIBE is a suite of serious games aimed at 14 to 19 year olds. Its modules cover subjects such as Construction and the Built Environment, Creative and Media, Society Health and

Development, Engineering and IT. Users navigate around a 3D environment in which they can carry out tasks and access media. One example of a task is bricklaying where the user learns how to lay bricks using different bonds, etc. This enables a user to practice in a non-hazardous environment without wasting resources, therefore, making it eco-friendly (SSAT, 2010).

2.3.3 Serious games and virtual characters

Although serious games offer the user free movement and new ways to explore and navigate learning material, they can lack something such as direction and structure (Ieronutti and Chittaro, 2007). In fully interactive virtual worlds, where the user explores the environment freely, the danger is that the user may wander around and explore but not complete tasks. This problem can be solved with the use of virtual agents or avatars (Sims, 2007) acting as guides or storytellers, prompting the user or giving instructions and taking them through a process towards a learning outcome.

Serious games have also been used in some innovative ways to help young adults with physical and intellectual disabilities to become more active (Yalon-Chamovitz and Weiss, 2008). The use of virtual reality technology can open up a whole new range of physical activities in a safe environment. Research into serious games is trying to help children on the autistic spectrum who can have problems with their emotional development (Jain et al., 2012, Bekele et al., 2013, Orvalho et al., 2010). They find it hard to learn the meaning of facial expressions and the emotions they portray. Serious games have been used to help teach these children starting with expression recognition and mimicking, followed by a game based on picking the correct expressions for different situations (Abirached et al., 2011).

Participants in this experiment expressed an interest in being able to pick different types of avatar such as humans, animals and aliens. Using an authored abstract avatar representation, would it be possible to hold a child's interest for longer? If a teddy bear representation was used, for example, would it result in longer play? One of the issues connected to this, however, is whether the same emotional content is portrayed when using an abstract avatar?

In a ten year review of educational virtual environments, Mikropoulos et al. (Mikropoulos and Natsis, 2011) suggest in their conclusion that serious games are becoming more commonplace in the classroom and, as a result, more research is needed into their ability to help in the learning process. In addition, they suggest avatars are another area that requires more study. This thesis' author would agree with both these points. Virtual humans within VEs are now quite popular. However, the level of visual and, particularly, behavioural realism that they portray still falls short of what is observed in real life.

2.3.4 Summary

Serious games are an interesting new medium offering often novel approaches to sometimes difficult to teach subjects. Such difficulties can be the result of expense or tasks being too dangerous to approach without prior training (Ieronutti and Chittaro, 2007, Falloon, 2010, Dalgarno and Lee, 2010). Virtual characters often act as guides or storytellers within these environments, and are used to steer users towards a learning outcome (Sims, 2007) or they act as avatars to represent the user in an immersive virtual environments (Vinayagamoorthy et al., 2004). VR and serious gaming can be used to train medical professionals (Graafland et al., 2012), virtually teach police officers to conduct interviews (Frank et al., 2002) or

allow remote collaboration on tasks (Roberts et al., 2015b) in a manner more akin to the real world.

Commercial gaming titles command large budgets, much larger than those of serious games. However, it is perhaps within serious games/VR applications that higher quality representations of virtual character NVB would be of most benefit. Using emergency planning as an example, is it not essential to convey as much behavioural detail as possible? Facial expressions, subtle behaviour and realistic full body movements convey a lot of information. The look of shock on a person's face may well appear before they shout "look out!" or "move!" In real life these cues would be valuable, so should they not be reproduced in serious games? The combination of new knowledge about the learning capabilities of serious games and creating more realistic virtual characters will not only help justify their use, but also enrich and expand their capabilities. Perfect realism may not be a necessity in order to achieve optimal effectiveness. It is possible that too much detail could distract or offer no improvement to learning outcomes. Studying the balance of visual and behavioural realism to enable the identification and interpretation of behaviours would contribute towards more effective virtual characters and serious gaming applications.

2.4 Chimpanzee behaviour and virtual animals

2.4.1 Introduction to Chimpanzee behaviour

When people visit zoos, they often look at the animals and perhaps laugh and gesture in amusement at them as they perform behaviour that people may find amusing, but what do these behaviours mean? What does the human response to the behaviour mean to the animal?

The visual displays of a Chimpanzee (*Pan Troglodyte*) to a watching visitor can be entertaining. However, the Chimpanzee may be responding to the body language and posture of that same visitor who may, unknowingly, be displaying a posture that the Chimpanzee finds threatening. A human facial expression can mean something completely different to a Chimpanzee than it does to another human. Chimpanzees and humans have similar underlying musculature to enable facial expressions (Vick et al., 2007) but their meanings are different. Research is ongoing looking at the evolution of facial expressions across primate species (Parr et al., 2007). By simulating the behaviour of these animals and allowing a user to interact with them in a VE, would it be possible to educate people on how best to view the animals with minimal impact on their daily routine? This would hopefully allow the animals to react in a more natural manner which would result in the visitor getting a truer representation of real animal behaviour. Visitors would also be able to identify why an animal was displaying particular behaviour. If a visitor was then able to identify the behaviour, the VE would have contributed to viewer education, which is one of the aims of a zoo (WAZA, 2005).

Chimpanzees are our closest animal relative. Indeed, their intelligence can be compared to that of a young human child (Tomasello et al., 2003b, Premack and Premack, 1994). They are fascinating, intelligent creatures that are capable of many varied behaviours, but how much does the average person really know or understand about them? Most people have seen Chimpanzees in zoos where they may spend ten minutes observing them, or on television in a documentary which tends to be about zoo life. In the worst case scenario, a person's only exposure to these beautiful creatures may be television commercials wherein they have been trained to act out human behaviour. One of the main aims of a modern zoo is conservation education (WAZA, 2005) thus how do we enhance an average person's visit

to the zoo to enable them to gain a far better understanding of the amazing array of creatures they have come to see? The first part of this section concentrates on primates, in particular on Chimpanzees. However, this could be extended to other species in the future. With the use of VEs it is possible to create virtual animals that could aid and enhance conservation education. A user, by interacting with the virtual animal, could build up an empathy for them and may want to learn more about them, for example such interaction may hopefully raise awareness of any perilous situation that the species may find itself in in the wild, through no fault of its own.

2.4.1.2 Chimpanzee hierarchy

Chimpanzees are social animals and live in hierarchically organised groups. At the top of this hierarchy is the alpha male, the highest ranking male in the group. As the alpha male he will have access to more food and females and, as a result, have higher reproductive success, as supported by the priority of access model (Murray et al., 2007). Beneath the alpha male are the rest of the group, from high ranking males and females to low ranking males and females. A Chimpanzee's ranking has many implications, for example, a high ranking or dominant female will have to travel less for food. Her foraging area will be smaller and more plentiful. Her subordinates, however, will have much larger foraging areas and they will be less abundant with food (Murray et al., 2007).

The effect of rank on food availability is further illustrated in several studies (Call, 2001, Tomasello et al., 2003a, Hare et al., 2000). These papers placed a dominant and a subordinate Chimpanzee into a controlled experiment environment. The two Chimpanzees are placed into opposite side rooms and can see out into a main room. Food is then placed

in the main room and the Chimpanzees are released. Both Chimpanzees know where the food is and the dominant Chimpanzee takes the food. The experiment is then re-run but this time placing visual obstacles in the dominant Chimpanzee's view. In this situation the subordinate has more success. Finally, the experiment is run with some food visible to both Chimpanzees and some only visible to the subordinate. The subordinate is then given a slight head start. In this case, the subordinate takes the food that only it can see. This illustrates two points: firstly, dominant individuals have better access to food sources. Secondly, Chimpanzees are capable of following eye gaze (Call, 2001, Tomasello et al., 2003a, Tomasello et al., 2001, Tomasello et al., 1999, Hare et al., 2000). In other words, they know what the other is looking at. It also shows they are capable of planning ahead to give them the best chance of success.

Male Chimpanzees form coalitions to strengthen and improve their position or rank in the group. Together they are stronger. Coalitions can be identified by behaviour such as associations, grooming, proximity, patrols and meat sharing (Mitani et al., 2002). Mitani and Watts (Mitani and Watts, 2001) report incidents of oestrous females begging for meat gained from hunts and being ignored. Yet the meat was shared between males reciprocally, between specific individuals. The author's study suggests that hunting for meat is an activity that improves male social relationships, that is, strengthening coalitions. Research suggests that these coalitions or bonds are long term. Mitani (Mitani, 2009) uses grooming and its frequency to identify these bonds. It was found that bonds were longer lasting when grooming was equal between males.

Grooming is an important behaviour in primates. It can reinforce relationships and bring calm after disputation, or help avoid confrontation. Puga-Gonzalez et al. (Puga-Gonzalez et

al., 2009) conducted a study that looked at confrontation and meetings between Macaques. They developed a computer simulation which decided whether the Macaques, on meeting, would have a confrontation or whether one of them will groom the other in order to try and avoid it. Confrontations in primates are extremely dangerous and can result in serious injury or even death (Goossens et al., 2005). In an effort to avoid these confrontations in primates, some zoos design enclosures to encapsulate escape routes and visual barriers, that allow for fast keeper intervention (Coe et al., 2009). The design of enclosures is another area where the use of VEs and virtual characters could be utilised (Stricklin et al., 1995).

2.4.1.3 Mimicking Behaviour

Experiments have been conducted looking at the effects of mimicking Chimpanzee behaviour when caring for them (Jenvoid, 2008). The results showed that the Chimpanzees displayed more playful behaviour when interacted with in this way as opposed to normal human interactions. Jenvoid (Jenvoid, 2008) suggests that keepers should adopt this approach when caring for Chimpanzees. Another study (Case et al., 2015) examined the use of Chimpanzee food calls on group movement behaviour at feeding times, finding them slightly more effective than using human communication. If this approach became more widely adopted, a virtual simulator could be a good way to teach the behaviour, enabling animals to be seen from all angles and distances with spatial audio. In addition, scenarios could be repeated, giving VR several advantages over the use of fixed viewpoint video footage. This would enable students to be taught approaches and to learn more about the reasons behind any behaviour before they were placed in a zoo. Zoo animal enclosures are often built to mimic the wild (Melfi, 2009) in an effort to stimulate wild behaviours, so why not take the same approach when interacting with the animals?

2.4.2 Zoo visitor education

Conservation education is one of the main aims of zoos (WAZA, 2005). One of the obstacles that zoos have is that most people visit as a form of entertainment (Ross and Gillespie, 2009). How do you then educate and inform visitors in an entertaining way? Currently, visitors learn from viewing animal exhibits and reading signage. In some instances they receive talks or interpretive presentations about the animals. Studies have been conducted to compare the knowledge gained between interpretive and fact only presentations (Visscher et al., 2009). These studies showed that people retain more information from interpretive presentations. What if it were possible to go one step further and engage the viewer in a VE, where they could view the environment from the animal's perspective? What if VE could be used to teach the viewer about the animal's behaviour and the motivations behind it? This educational resource could also be used to educate staff about animal behaviour and possibly help them avoid or diffuse dangerous situations before they happen. VR simulators have often been used to teach or simulate processes that would otherwise have been very dangerous or too complex to teach, such as emergency planning.

2.4.3 Virtual animals

In order to learn about animals people have traditionally visited zoos, museums or natural habitats. A new medium is now available to us, the "Electronic Zoo" (Davies, 2000). Virtual environments allow the modelling of both animals and their environment, thus allowing the study or modelling of virtual ecosystems, individual species behaviour, or even the enclosure design of captive animals. Heleno and dos Santos (Heleno and dos Santos, 1998) looked at modelling a virtual ecosystem of an estuary of the River Sado. The ecosystem contains mullets, dolphins and jelly fish. All of the animals have the same underlying

architecture. However, some are drawn to each other, such as predators, and some are repelled as prey. The surrounding environment can also affect the ecosystem, using for example a pollution variable to affect animal health. In a species specific study, Tomlinson et al. simulated the emotions and behaviour of gray wolves (Tomlinson and Blumberg, 2001) with the aim of enabling users to better understand a species relationships with its environment. Stricklin et al. (Stricklin et al., 1995) used animats (virtual animals) to investigate enclosure design. They undertook this using a two dimensional approach. The enclosure is modelled and the animats are monitored to record space usage. The outputs included distance moved, direction changes, boundary collisions and personal space violations. The author suggests that this technology has potential uses for animal welfare.

2.4.3.1 Virtual Pets

Virtual animals and creatures have become commonplace in games from Tamagotchi pets (Bandai, 2016) to Nintendogs (Nintendo). In typical examples, the game player must feed the animal and play with it. Failure to look after the animal correctly will result in its demise (Fogg, 2002). Kaplan (Kaplan, 2000) looks at the role of uselessness in the design of virtual pets and suggests that one of the intriguing design features of virtual pets is that they are useless and have no practical application, yet people of all age groups spend time playing with them, perhaps suggesting that users feel a duty of care for them. Other research illustrates the use of virtual pets to improve mirror neuron function in autistic children (Altschuler, 2008) and to help children with cognitive impairments who have difficulties with interpersonal relationships (Marti et al., 2005). The use of virtual pets allows most children to develop emotions towards them (Chen et al., 2011, Kusahara, 2001) and studies have examined how emotional attachment to virtual pets can be used to encourage

behavioural change (Dillahunst et al., 2008). Using virtual pets might enable zoos to teach visitors about different species in a more entertaining way, allowing them to continue learning after the visit. This would, therefore, contribute to conservation education.

2.4.3.2 Virtual Primates

Piccault and Collinot (Picault and Collinot, 1998) investigated the creation of multi agent systems by simulating primate societies. They suggested that there are links between primate behaviour and multi agent systems, by simulating a Macaque group of between five and thirty individuals. They used a reactive model to simulate social concepts and to look at distinctions between organisational levels and they split the organisational levels into three using the Cassiopeia method. Level one included domain dependant roles such as threat, submit and eat. Level two looked at relational roles such as affiliation, dominance and kinship. Level three considered organisational roles such as coalitions and alliances. The study then went on to look at social cognition using the pheromone model. In this model, when an event occurs, a social pheromone is given off into the environment which, when perceived by an animal, causes a change in its relational and organisational behaviours and knowledge. Piccault and Collinot then looked at simulating attention using distance and field of view. They ran a test with a population of between five and thirty agent animals looking at dominance behaviour and found that a linear hierarchy would form, aggression would go down and coalition go up. This would then collapse and the process start again. In a separate study examining Chimpanzee behaviour using artificial intelligence (AI), Chimpanzees search for food and mates resulting in a group hierarchy being simulated, with males demonstrating grouping behaviour and females being more solitary (Irenaeus et al., 1994).

In *Gorillas in Bits* (Crowell, 1997) the user becomes an adolescent gorilla and part of a gorilla family. An adult male and female in the habitat are both resting. The user can approach the adults. However, there are rules. The male has a large personal space and intrusion into this in an aggressive way will result in behaviour such as chest beating, charges, vocalisations or even physical attack. The female gorilla is easier to approach and will allow the user to approach her in a slow submissive way. If she does not respond then the user can groom her. However, an incorrect approach to the female gorilla will also result in an aggressive behaviour. One nice feature of this simulation is the “Time Out” function. The authors knew that if children had the opportunity to invoke an aggressive response they may, in some cases, deliberately do so. If they did this the screen would turn black and a message would come up telling them that they were in “Time Out”. This is an excellent idea as the simulator is teaching users about gorillas and their behaviour, not how to annoy them. This approach would be sensible to adopt in the current project if the simulation was to be used by the general public. Hay et al.’s (Hay et al., 2000) aim was to help children gain a greater understanding of the biomechanics of gorilla behaviour/movements, such as grooming and knuckle walking. The children in the study also used the Virtual Gorilla simulator at Zoo Atlanta. They utilised a HMD, were inside a gorilla habitat, and assumed the role of an adolescent gorilla. After using the simulator they made models of gorillas in poses using pipe cleaners. They also wrote down some instructions for behavioural rules that were input into the simulator and could be viewed by them to see what had occurred. The study demonstrates how VR can be used to teach young children on visits to zoos.

2.4.4 Summary

To summarise, the literature informs us that zoos have an obligation to teach visitors about their animals and exhibits (WAZA, 2005) and need new and innovative ways of doing this (Davies, 2000). The literature identifies that certain animals respond positively when keepers care for them by mimicking the animal's behaviour (Jenvoid, 2008, Case et al., 2015). Some research has been carried out regarding teaching zoo visitors about primate behaviour (Hay et al., 2000) (Crowell, 1997) while other studies have successfully used AI to simulate primate society (Picault and Collinot, 1998) (Irenaeus et al., 1994). The literature also shows that people can develop emotional connections with virtual pets (Chen et al., 2011, Kusahara, 2001) and that this connection can be a good tool for engaging people with bigger issues (Dillahunt et al., 2008). All of the above suggests that VR and serious gaming could be a good way of facilitating learning for students, zoo staff and visitors. However, before moving onto discussing more advanced studies, it is important to understand whether CGI used in VEs contains sufficient detail for the identification of animal behaviour. By isolating the visual variable and scripting the user experience, it is possible to ensure that users see the appropriate learning material. A comparison can then be made with fixed viewpoint film, before moving onto unscripted user controlled interactions and the added benefits or distractions that they bring in later studies. In this, and subsequent studies, the researcher aims to build on these foundations and show evidence in support of using serious games and virtual animals as a new and exciting way to educate people on animal behaviour, whilst ensuring that 3D technologies are not used simply for the sake of using them.

2.5 Virtual humans and human behaviour

In computer games and virtual environments, characters are used to represent users or players and non-player characters. Characters are created with various levels of detail depending on the application and the platform. They can be hand authored or automatically generated using 3D reconstruction. Within immersive collaborative virtual environments, the standard approach captures motion data from the user to drive an authored virtual character constructed of separate body parts in real-time. Approaches vary from simple head and hand tracking to systems which include eye gaze (Vinayagamoorthy et al., 2004). Studies have demonstrated how a number of non-verbal behaviours can be shown using this approach (Slater et al., 2013) while this technique has also been successful for conducting collaborative tasks (Roberts et al., 2003).

The simultaneous capture of full body motion and more subtle facial expressions is more challenging. Typically, cameras are positioned around a capture area for full body motion capture and a separate camera is mounted close to the face for capturing facial expressions. This technique works well for recording high quality animation, but has limitations if used for real time collaborative communication due to markers causing discomfort for users and the setup time required. Affordable commercial software, such as Faceshift (Faceshift), allows for real time marker-less capture of facial expressions. However, this approach still requires a depth based camera to be positioned close to the face. One possible solution for collaborative communication, where both full body and subtle non-verbal communication are important, is video based 3D reconstruction. This method requires no markers nor any other tracking technology, allowing unencumbered movement while supporting facial expressions (Roberts et al., 2015a), eye gaze (Roberts et al., 2013b) and, with sufficient

frame rate, micro expressions. Micro expressions are facial expressions that occur in a fraction of a second (Ekman and Friesen, 1969) which provide clues to our emotions.

One current challenge faced by character developers is how to incorporate more realism into the authored characters. When examining virtual characters, two measures used are realism and believability. Of particular interest to this study is visual realism, predominantly appearance and behaviour, and, in turn, our perception of it (believability). When a user encounters a virtual alien, they have no real knowledge of its behaviour or appearance handing full creative freedom to talented artists, animators and storytellers. However, if a character is human or human-like a problem arises in that we, as humans, view and study each other every single day. We know exactly how we should look, move and sound. Thus, flaws in virtual humans will be spotted instantly.

The theory of the Uncanny Valley, created by Japanese roboticist Dr Masahiro Mori (Bryant, 2005, Groom et al., 2009), suggests that there is a correlation between user perception and realism. Once a character gets close to a realistic appearance, but not quite, it can become somehow sinister, or less believable, than a much simpler cartoon version of itself. This is evident when watching movies with computer generated people in them (Gutierrez, 2011) as uncanniness impacts upon experience. Because to this, some film makers have made characters less human to avoid the ‘uncanny valley’ (Brenton et al., 2005, Weschler, 2002). Similarly, in computer game development, characters have been produced with less visual realism in an effort to overcome uncanniness using techniques such as toon rendering (Brenton et al., 2007).

Developers might be guilty of ignoring uncanniness in their pursuit of visual realism and the higher level of presence this is thought to yield (Tinwell et al., 2011). Interestingly, whereas one may strive for higher levels of visual realism in environment modelling to enhance levels of presence, as illustrated by (Khanna et al., 2006, Slater et al., 2009), the theory of the uncanny valley suggests that higher visual realism in virtual characters could cause developers a problem. Pausch et al. report a break in presence if characters do not respond to users within a virtual environment (Pausch et al., 1996). A break in presence (BIP) is when focus switches from the virtual environment to the real (Slater and Steed, 2000, Slater et al., 2003) illustrating that both visual and behavioural realism are important for supporting presence. Tromp et al. suggest that, if quality of visual appearance and behaviour do not match, then characters cause unease (Brenton et al., 2005, Tromp et al., 1998). However, Thompson indicates that, over time, players may become used to the effect of the uncanny valley (Thompson, 2004) and may now accept slightly creepier characters.

Our perception of photo realism, the researcher suggests, is a moving variable that changes according to what users consider to be at the technical forefront at the time of viewing. This is demonstrated by Vinayagamoorthy et al. (Vinayagamoorthy et al., 2006) who, in 2006, used a character from the 2004 game Half Life 2 as an example of highly photorealistic rendering. The author of this thesis suggests that this is no longer the case (see *Fig. 6*) as advances in materials, texturing and lighting now afford a more realistic rendering of skin and eyes (see the Activision Blizzard character), leaving the older Half Life character looking flat. It is evident that designers of synthetic characters are closing in on visual realism and perhaps more attention should now be given to balancing behavioural realism with it.



Fig. 6. Left image illustrates Virtual human by Activision Blizzard 2012

(<http://www.iryoku.com/blog>), the right image is from Valves Half Life 2 from 2004

(<http://half-life.wikia.com>).

2.5.1 Types of virtual character

Virtual characters can be split into two categories, avatars and agents. An avatar is a character whose behaviour is controlled by a human user and an agent is a character whose behaviour is controlled by the computer using artificial intelligence (Nowak and Biocca, 2003).

Avatars

Hand authored avatars can take any form. Here the focus is on the human form and different methods of creating/representing it. Hand authored avatars are created using modelling (Autodesk, 2015a, Autodesk, 2015b, Blender, 2015) and texturing packages (Adobe, 2015). In immersive collaborative virtual environments they have traditionally been an unskinned character comprising of body parts transformed and rotated by the real time motion tracking of a user. An example of this type of avatar is used by Roberts et al. (Roberts et al., 2003).

In this study, the term abstract avatar refers to a modelled and rigged CGI avatar. This means that the avatar body is one whole piece of geometry rather than individual body parts. Bones are then applied to this geometry in a process called rigging. When these bones are moved, the geometry of the character deforms accordingly, producing a more realistic appearance than an ICVE avatar. Finally, 3D reconstruction can be used to generate avatars automatically. For this study, the term video based 3D reconstructed avatar refers to a 3D avatar created in real time from 2D video streams, sometimes called a virtuality avatar (Roberts et al., 2013a). This medium attempts to be an exact replica of the captured user. This type of avatar is utilised in the study by Roberts et al. (Roberts et al., 2015a).

Agents

An agent is a computer controlled character that has a set of actions which can be triggered as a result of stimuli/input in the game world around them. For example, if an agent is set to a hungry state it may search for food. Once the need for food is removed, the controller (the agent's brain) may change the agent's state to carry out another function such as tracking another player. In games these characters are also referred to as a non-player character (NPC) (Millington and Funge, 2009). Artificial Intelligence (AI) is used to create these agent behaviours. "Artificial intelligence is about making computers able to perform the thinking tasks that humans and animals are capable of" (Millington and Funge, 2009). It is because of AI that the game industry is capable of producing NPCs that react to a player's every move in a given situation. This makes the gaming experience much more rewarding. Many of the techniques used in gaming AI would translate well for edutainment programmes. In most games the characters do not need to have an overly complex AI architecture and thousands of possible behaviours. They do, however, need to be able to

react correctly in a given situation. In entertainment titles, behaviours may be limited to those needed for task completion; however, if social interactions are required, then more complex AI will be needed. Additionally, the characters may need behaviour to adapt over time, for example, if a NPC in a game reacted in the same way to the same move every time, it would be unrealistic. Agents need to adapt to allow for more believable responses, even simple responses, which can produce powerful effects; for example, participants in one study (Philipp et al., 2012) displayed more frowning and smiling expressions when shown negative and positive stimuli provided by virtual agents. This illustrates the powerful effects that agents can have on user emotion if used well. Poznanzki and Thagard (Poznanski* and Thagard, 2005) examined the modelling of personality and behaviour. They recognised the growing need for virtual agents in fields such as electronic entertainment and other areas of human computer interaction. Poznanzki and Thagard used neural networks to model personality and personality change. The virtual characters are placed in different situations and their personalities change and adapt over time due to this. In Poznanzki and Thagard's study, behaviour is based on three things. The first factor is the likelihood of their personality to use such behaviour, secondly their mood and, finally, the situation that the character finds itself in. Each of these variables has a value assigned to it. These values can be increased or decreased over time dependent on the situations encountered. The resultant output behaviour is different depending on the value of the variables, thus the characters' personality evolves. The variable value is capped so that the characters' personality cannot be changed too much. The authors suggest that modelling character personalities in this way created behaviours that change over time in a viable psychological manner. Due to agents being computer controlled, they can be created as being logical and unbiased, making them useful for situations where human emotion may become an issue, an example being

police lineups (Daugherty et al., 2007, Daugherty et al., 2008). An officer may unintentionally display biased behaviour, whereas an agent will not.

2.5.2 Human behaviour

Human behaviour is both verbal and non-verbal. This section of the literature review outlines aspects of human behaviour that are desirable for virtual characters to be capable of demonstrating in social communication.

Non-Verbal Communication – NVC is communication using body language. It includes gestures, facial expressions, touch and eye gaze, and can even be affected by a person's visual appearance, such as clothing or their choice of hairstyle. It is powerful because it can not only shape our view of an individual within seconds of meeting them (Hinde, 1972) (for example, perhaps posture and eye contact giving away a lack of confidence) but also gives clues of intent which can be read; for example, an angry person walking quickly towards someone would immediately put the person being approached on guard.

Gross Non-verbal behaviour – The term gross non-verbal behaviour will be used to describe large full body motions such as walking, running or sitting.

Eye Gaze – Eyes play a large role in NVC. For example, people may glance at someone they find attractive but not stare. If someone has done something wrong they may look at the floor. Before boxing matches, pugilists often have a 'stare down' to try and intimidate their opponent. These examples are accompanied by other body language, however, without the correct eye gaze the meaning may change. Research has been conducted demonstrating

virtual characters' ability to accurately display gaze using both authored and video based 3D reconstructed avatars (Murray et al., 2009, Garau et al., 2003, Roberts et al., 2013b) with both types of avatar successfully doing so.

Proximity/interpersonal distance – This relates to the distance between two people who are communicating. Social rules on interpersonal distance and eye gaze have been shown to be present between users and the characters in virtual environments (Yee et al., 2007). Furthermore, users show signs of heightened physiological response as virtual characters approach them (Llobera et al., 2010). Interpersonal distance between an authored avatar and virtuality avatar is shown in Figure 7.

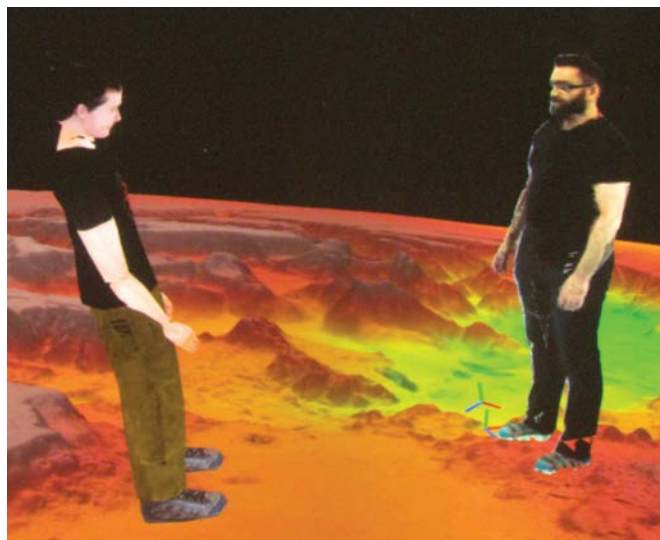


Fig. 7. Interpersonal distance between authored avatar and a virtuality avatar (Fairchild et al., 2015)

Orientation – Orientation is the angle at which an individual faces during communication.

Facial Expressions – Facial expressions can portray numerous emotions such as joy, surprise, contempt, sadness, anger, disgust and fear (Matsumoto and Hwang, 2011). It is,

therefore, unsurprising to know that they play a vital role in face-to-face communication. Nasoz and Lissotti (Nasoz and Lisetti, 2006) state that human to human interaction should form the basis for human to computer interaction (HCI). They explain that studies in the field of HCI have suggested that having an avatar with the ability to display facial expressions as part of an interface improves the user's performance. Nasoz and Lissotti's aim was to enhance HCI with the use of avatars and facial expressions. Firstly, they looked at capturing the user's emotional state. Once they had this information they could use it to aid the user. For example, if the user was worried, or stuck on a particular problem, the system could display expressions that would give encouragement or could calm the person. This might then allow the user to relax and focus on the task at hand, and generally make them feel at ease. It is possible to code facial muscle movements using the facial action coding system (FACS) (Ekman and Friesen, 1978) and skilled coders can use this to identify emotion. FACS is currently used to code muscle movements but could, in addition, be used to measure the quality of facial expressions in abstract and video based 3D reconstructed avatars.

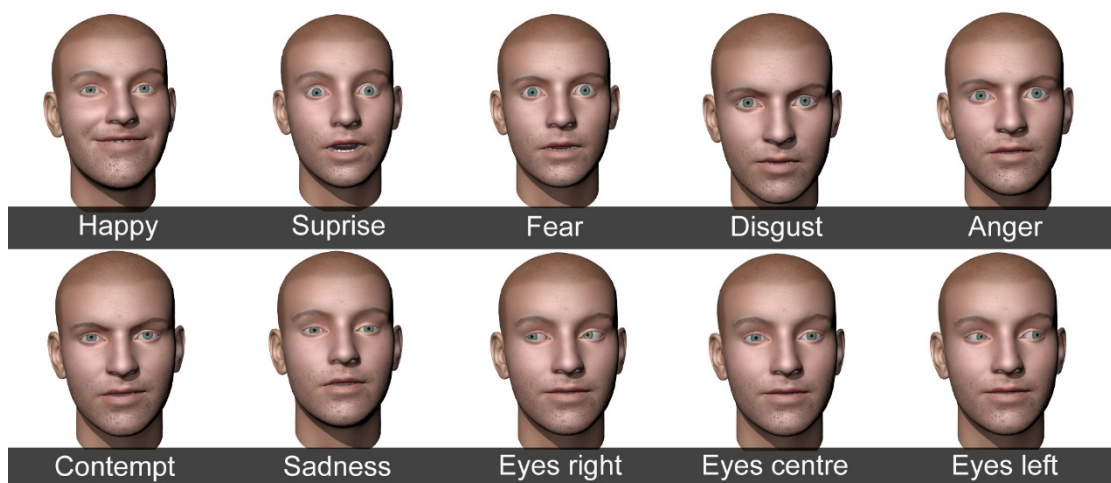


Fig. 8. Facial expressions and eye gaze portrayed by an abstract 3D avatar.



Fig. 9. Universal facial expressions of emotion and eye gaze. This figure shows a 3D reconstruction of the seven universal emotions and the quality of eye gaze in three directions.(Fairchild et al., 2015)

Figure 8 illustrates facial expressions and eye gaze on an abstract avatar captured using Faceshift (Faceshift) while Figure 9 shows the quality and clarity of facial expressions on a video based reconstructed avatar achievable with a good camera calibration. Both images illustrate the seven universal emotions (Ekman and Matsumoto, 2008). It should be noted that for 3D reconstruction, camera rig height can have an impact on reconstruction quality inducing a droop effect that can make a user appear sad, aged or unwell (Roberts et al., 2015a). This becomes worse as the user approaches the outer limits of the capture space. The camera setup for Figure 9 appears to be correct for the capture subject.

Micro Expressions – A micro expression is a facial expression that only occurs for a fraction of a second which gives away clues to emotion. They can be used in deception studies to identify leaked emotions (Ekman and Friesen, 1969). Endres and Laidlaw

examined the use of micro expressions to help medical students gauge how a patient was feeling during consultation (Endres and Laidlaw, 2009), whether a patient leaked a clue of feeling fear or sadness, for instance. These clues could then allow the medical professional to further question the patient regarding the situation.

Lip Sync – In virtual environments, or in computer gaming when characters speak, lip sync can be displayed alongside text on screen or a sound file. If sound files are played, then they should be synchronised with the character’s lip movements, facial animation and other gestures.

Bodily contact – Examples of bodily contact include, but are not limited to, ‘high fives’, handshakes on meeting, a hand on the shoulder to console somebody, a pat on the back for a job well done, or hugging to greet someone.

Posture – Posture can convey confidence, can show aggression or fear, happiness and sadness. An avatar with its shoulders back and chest puffed out, for example, would appear confident.

Head nods and shakes – A nod or shake of the head can illustrate agreement or disagreement.

Gestures – Some examples of gestures are pointing, beckoning and waving goodbye or hello (Figure 10). They can be used to show anger (such as by the shaking of a fist), more unpleasant finger gestures, confusion or lack of understanding (for example, by a shrug).

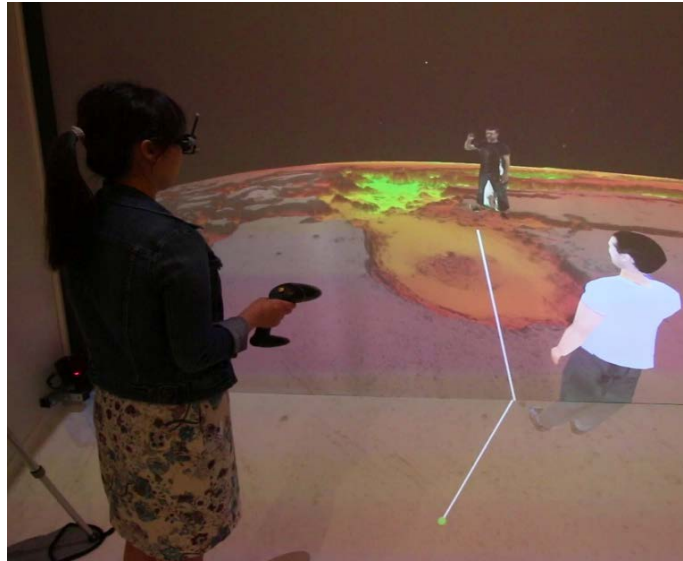


Fig. 10. Virtuality avatar waving. (Fairchild et al., 2015)

2.5.3 Challenges of representing non-verbal behaviour using a virtual character

If humans can form a view of a person within seconds based on NVC (Hinde, 1972), then it is surely important to attempt to accurately represent elements of this within virtual characters. Without such elements, much meaning is lost. This will be challenging as even blockbuster game titles such as *Skyrim* (Bethesda, 2011) lack subtle NVC. Specific applications require characters to display varying levels of behavioural realism. This combination of visual and behavioural realism, matching application, is what the researcher would term “relevant realism”. A first person shooting game such as *Call of Duty* (Activision, 2010), for example, might require high visual fidelity but does not need behavioural detail, such as facial expressions. However, an application to help autistic children learn about emotion (Jain et al., 2012, Bekele et al., 2013) may need less visual realism and more behavioural realism. If higher levels of behavioural detail are necessary to convey more subtle movements (such as facial expressions and eye gaze) developers need to ensure that avatars are capable of displaying them. Standard gaming avatars generally

allow a user to walk around an environment with limited NVB displayed. Applications requiring social interactions between users could currently be missing out on the majority of non-verbal interaction. If NVB is not to be lost this, technically, creates an interesting challenge. If authored abstract avatars are to be used, then tracking hardware needs to be placed close enough to the face to identify facial muscle movement (Faceshift, 2015). This is adequate if the person is sat in a chair, but not if he/she is moving around. The tracking of full body motion and facial animation simultaneously can be achieved, but requires more complex tracking hardware (Faceware, 2015). Video based 3D reconstructed avatars (Fuchs et al., 1994, Raskar et al., 1998) may offer a solution to this, in that the user can be captured and represented faithfully in 3D, including eye gaze (Roberts et al., 2013a) and facial expressions (Roberts et al., 2015a) while allowing free viewpoint navigation around him/her, without the user being covered in tracking markers or wearing a head mounted facial tracking camera, thus offering more natural unencumbered movement.

If 3D reconstructed video is a potential solution, it also has its own set of challenges. For example, Beck et al. (Beck et al., 2013) used 3D reconstruction for group interactions. All the participants wore 3D shutter glasses to enable stereoscopic viewing, therefore losing the ability to display eye gaze and facial muscles around the eyes. The use of these glasses counteracts the proposed benefits of 3D reconstruction. This problem is illustrated in Figure 11 which is taken from a study by Fairchild et al. (Fairchild et al., 2015).



Fig. 11. This figure demonstrates the problem of occlusion (for the viewer) when using various display devices. Left: no stereo (full facial expressions); centre: stereo glasses (eye gaze and some facial features obscured) and right: HMD (most of the face obscured, identification of facial expressions not possible).(Fairchild et al., 2015)

One challenge for future developers may be the lack of stereo for such applications, if two-way communication is required. Of course, even without stereo, 3D reconstruction still offers free viewpoint navigation (Roberts et al., 2015a). This is demonstrated in Figure 12. Additionally, most VR motion capture/tracking suites mount markers on the glasses used for stereo and on the body to allow for user interaction (Vicon, 2015). These markers would also be reproduced in the capture process. If these traditional systems were no longer to be used, a user skeleton may well have to be derived from the reconstruction in order to allow for collisions and the general object interactions that users are now accustomed to in VR. This can be achieved using depth-based approaches, which can even support multiple participants (Kinect).

Another potential challenge for 3D reconstruction is the increasing popularity of HMDs for home use, gaming and research. These present challenges to both authored abstract and 3D

reconstructed avatars. The current size of these units results in most of the face being occluded, thus making capture of facial expressions by either 3D reconstruction or tracking (marker or markerless) unachievable. Of course, gross full body tracking is still possible using combinations of hardware, such as the forthcoming consumer Oculus Rift (Oculus) with tracked controllers and the OMNI 360 degree treadmill (Virtuix, 2015). The lack of facial tracking may be acceptable for standard gaming applications but, in serious contexts where NVB is necessary, it is not acceptable, perhaps therefore pointing to certain types of hardware/software combinations being more suitable for specific applications.

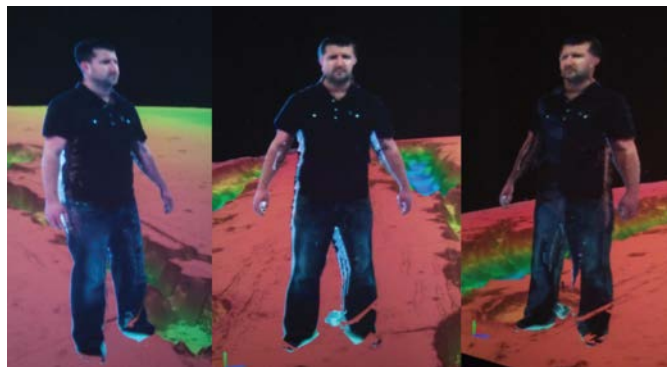





Fig. 12 Depiction as to how the 3D reconstructed user is viewed from different viewpoints as one user moves around him. This free viewpoint navigation offers the user interactive functionality over and above that offered by traditional video conferencing, allowing users to view NVB from any angle (Fairchild et al., 2015)

Device	CAVE/OCTAVE	HMD	Power Wall	Desktop PC
~ Visual Immersion	270 H 200 V for 3 wall and floor CAVE. 360 H 200 V for OCTAVE	360 horizontal 360 vertical	90 horizontal 90 vertical	90 horizontal 90 vertical
Stereoscopic	Yes, requires head tracking for viewpoint update.	Yes, requires head tracking for parallax	Yes, requires head tracking for parallax	Yes, requires head tracking for parallax
Capturing of gross full body	Yes, given tracking system.	Head tracking in built. Yes, given tracking system.	Yes, given tracking system.	Upper body, usually optical such as Kinect.
Capture of Facial expressions	Partial if using stereoscopic glasses. Full possible, however, authored avatar requires camera near face.	Partial, as part of face occluded by display.	Partial if using stereoscopic glasses. Full possible, however, authored avatar requires camera near face.	Full – as the user is seated in a fixed position it makes marker or markerless tracking easier to achieve.
3d Reconstruction	Yes, given sufficient space and correct camera setup.	Yes, but excluding upper face detail.	Yes, given sufficient space and correct camera setup.	Partial.
Size	Large	Small	Medium	Small
Cost	High	Low	Medium	Low

Table 2. Table suggesting that full 3d reconstruction requires space and. generally more expensive hardware

An interesting difference between authored avatars and 3D reconstructed avatars is that the former can be easily deployed as agents. Animation can be stored because it is known where each bone and vertex is positioned at all times and these can be reverted back to idle poses or blend between animations. However, because of the way video based 3D reconstructed avatars are created, the same procedure cannot easily be implemented on them. A possible solution for this would be to create a hybrid abstract/3D reconstructed character. A static version of the character mesh would first need to be 3D captured, holes filled, topology and geometry reworked and, finally, rigged before it could be deployed as an agent. However, it would be very challenging to blend fluid 3D reconstructed character animations with ever changing geometric and textural detail into a seamless agent without a lot of post processing.

	Hand authored avatars		Automatically generated avatars
	ICVE Avatar	Abstract Skinned Rigged Avatar	Video Based 3D Reconstructed Avatar
			
Gross full body behaviour	Yes, translation and rotation of body parts.	Yes, animation of rigged character.	Yes.

Subtle behaviour	Eye gaze and hand movement possible, texture changes could allow facial expressions.	Possible if either pre-recorded or an adequate tracking system available.	Yes, dependant on reconstruction quality and speed.
Proximity	Yes		
Eye gaze	Yes.		
Real time facial/micro expressions	Facial expressions Possible through texture changes; micro expressions would be problematic.	Yes.	Yes.
Truth Bias	Yes.	No.	Unknown.
Free viewpoint update	Yes.		
Suitable for an avatar	Yes.		
Suitable for an agent	Yes.	Yes.	No, because of mesh detail and texture mapping change, making transitions between animation sets problematic.
High realism possible	Hard to achieve with separate body parts.	Yes.	Yes.

Table 3. A comparison of virtual character attributes

Table 3 illustrates a breakdown of virtual character attributes. It shows that all types of avatar are able to display gross full body behaviours, interpersonal distance and eye gaze,

and can permit free viewpoint update, thus allowing a viewer to walk around them. However, when looking at more subtle behaviours, perhaps to represent facial expressions, for real time communication, the author of this thesis would suggest that ICVE avatars are not the best option. Instead, abstract or video based 3D reconstructed characters that are better equipped to display more subtle behaviour, such as facial expressions, would be better. Comparison with the literature suggest that avatars with a lack of facial expression display a truth bias higher than normal (Steptoe et al., 2010) whereas abstract avatars with facial movement (McDonnell et al., 2012) report truth bias in line with psychological literature (CF Jr Bond, 2006). It should be noted however that only a small number of studies could be identified, so no firm conclusion can be made. There are, as yet, no tests on whether 3D reconstructed characters cause truth bias. The most flexible character type is abstract, as it is capable of displaying gross and subtle behaviour and can be easily deployable as an agent or avatar. The video based 3D reconstructed avatar can be used as an agent but is not generally utilised as such due to complexity and the time associated with doing so.

Virtual characters are so complicated that people struggle to encapsulate all the facets that are needed to create the ideal character and, often, compromises must be made. These do not always impact on the role or effectiveness of the character. For example, not every character needs to speak, thus when designing a character it is possible to narrow down this specification. A careful selection of components means that it is possible to create a character that is still fit for purpose. It is also important to create a compelling character that displays a balanced level of detail in terms of visual appearance and behaviour so as not to fall prey to the uncanny valley (Bryant, 2005, Groom et al., 2009). Once factors, such as those discussed above, have been taken on board in order to create a meaningful and useful virtual character, then it is necessary to look into how and why the character will be deployed.

Will they be a guide, tutor or storyteller? Will they act as an opponent or collaborate with the user? How will they be controlled? Will they be an agent or an avatar based character? All these characteristics have their advantages. For example, having an agent controlled virtual character to interact with, means that a single user can learn or play in a virtual environment independently. However, an avatar could be controlled by an expert in a particular field. This would work particularly well for animal behaviour, as an expert would know when an animal should be displaying certain behaviour. Lim and Reeves' (Lim and Reeves, 2010) research into agents and avatars suggests that users have a higher sense of arousal when interacting with an avatar rather than an agent and this may, in turn, lead to a better attention span and improved learning. Using either technique for controlling virtual characters is an exciting way to help engage learners. Babu et al. used virtual characters to teach conversational protocols (Babu et al., 2007). These protocols, from South Indian culture, were both verbal and non-verbal. An experiment was conducted which contrasted learning via the use of the virtual characters with that of learning using an illustrative guide. The virtual characters displayed actions and then the user repeated them. This was recorded and tracked in real-time, so that a second virtual character representing the user could be viewed carrying out the actions. The results showed that learning via the virtual character gave better results. Variation between results was also four times lower in the virtual characters' group which suggests that the use of virtual characters was a more reliable teaching method in this instance, therefore, supporting the use of virtual characters as an effective way of teaching body language.

2.5.4 Effect of virtual characters on users

A study by Bailenson et al. (Bailenson et al., 2006b) suggests that the portrayal of behaviour in avatars is of importance not only for VEs but also for telecommunications, online gaming and any other media that require the user to be embodied as an avatar. While Gillies and Slater suggest that characters that respond to others with appropriate non-verbal behaviour can improve presence in virtual environments (Gillies and Slater, 2005). Previous studies have shown that users often prefer not to use accurate embodiments (Becerra and Stutts, 2008) but how does avatar representation impact upon interpersonal trust? Avatars are becoming more commonplace and are being used to represent users in virtual worlds, games, instant messaging, social media and more serious training applications. What effect does the selection of an avatar have on the user and others? Bailenson et al. state that, as avatar realism rises, so does the feeling of co-presence while simultaneously decreasing self-disclosure (Bailenson et al., 2006b). Video conferencing has achieved the lowest scores for verbal and non-verbal disclosure (self-disclosure being when an individual divulges information about themselves or their past experiences (Derlaga and Berg, 2013)). This would suggest that the more realistic an avatar becomes, the more guarded an individual becomes concerning what they disclose.

In a study by Rosenburg et al. (Rosenberg et al., 2013) participants assumed the role of either a super hero or a passenger in a helicopter. They were given one of two tasks, either helping a diabetic child find some insulin or having a virtual tour of the city. After the virtual environment experience had finished the experimenter knocked over a box of pens. Participants in the super hero condition helped pick up the pens significantly more times than those who opted to be the passenger in the helicopter. Peña et al. explored how an avatar's appearance impacts upon the wearer (Peña et al., 2012), examining changes in

behaviour. Participants were represented by either a professor or super model avatar. The study found that behaviour was influenced. Participants represented by the super model were more age conscious. Fox et al. (Fox et al., 2013) examined the effects of sexualised avatars on female participants. They described the participants as “wearing” the avatar. Participants were split into two groups: one wearing the sexualised avatar and one the non-sexualised avatar. The results found that the users of the sexualised avatar reported more body related thoughts. Fox et al. stated “Users of video games, online social worlds and other virtual environments should be made aware of the potential implications of the avatar they embody.” The researcher agrees with Fox et al. on this point, as evidence suggests that the avatar that one “wears” subconsciously affects our actions, temporarily altering our behaviour (Peña et al., 2012, Rosenberg et al., 2013, Fox et al., 2013) during and after VE exposure.

Users of virtual environments are not just affected by their own avatar. Pan and Slater examined the response of male participants to attractive female characters within an immersive virtual environment, with the results suggesting similar physiological, behavioural and subjective responses to reality (Pan and Slater, 2007). In a study by Behrend et al. (Behrend et al., 2012) avatar appearance was investigated, this time in the context of a job application. The study showed how attractive avatars were given more favourable feedback than less attractive ones. Interestingly, however, extremely attractive avatars could also produce a negative effect: “Non-linear effects of attractiveness are also possible, leading unusually attractive avatars to be viewed as unprofessional or dishonest. These dynamics are ripe for study”. This indicates that not only are the users of avatars influenced by their appearance, so are others. One problem may be that most avatars are generally quite simple representations of the wearer, as they do not portray the user’s NVB. Without NVB, verbal

communication and avatar appearance may well be the only information that someone has to inform them what the user behind the avatar is like. In some cases, users may not want others to know what they look like, with studies showing a correlation between body image and a user's desire to be someone else (Becerra and Stutts, 2008). Avatars can be a form of escapism. Utilising them, anyone can become a hero or a star. People may well pick an avatar that represents what they would like to be, aiming to be seen in the best light possible. An example would be a person who might use a profile picture of themselves from when they were ten years younger and super fit. In some cases, unfortunately, it is not possible for people to pick an avatar that resembles them as "the vast majority of games, both online and offline, do not allow for the creation of avatars with a non-white racial appearance." (Dietrich, 2013).

In another study that used avatars to look at obesity (Dean et al., 2009), it was found that participants interviewed by thinner characters stated that their own reported weight and body mass index (BMI) was correct. However, when interviewed by overweight characters participants said their own reported weight was incorrect and that they had a higher BMI. This suggests that participants were able to relate more to the overweight avatar and be more truthful, perhaps viewing them as less judgemental. This could represent a problem, as it illustrates that avatar appearance can be used to manipulate user interactions. This technique could also be used positively to educate users on health issues, such as in the doppelganger study (Fox and Bailenson, 2010). In this study an avatar representation of the participant was created that gained and lost weight according to participant actions.

Our perceptions of a virtual character can change depending on whether it is an agent or an avatar. More is expected of characters that are user controlled. Guadagno et al. (Guadagno

et al., 2011) found that a smile from an agent increased empathy whereas a smile from an avatar decreased empathy. They suggest that this may be the result of behavioural and visual realism not matching. The researcher agrees that expectations of agents are not as high. The researcher would suggest that most users would accept the fact that to replicate every human function would be a massive task, so are quite forgiving of agent shortcomings. However, when users know that they are interacting with a person their expectations may be higher.

More realistic avatar appearance and behaviour should facilitate more natural social interactions between users. However, transformed social interaction could pose a problem in the use of virtual characters (Bailenson and Beall, 2006, Bailenson et al., 2004). For example, it is possible to conduct real-time analysis of facial expressions (Yap et al., 2012), however, if unwanted emotions were then filtered out the meaning of a conversation could easily be changed. This kind of technology would make it very difficult for users to know when to trust authored avatars, perhaps making real-time video based 3D reconstruction approaches more relevant for communication.

2.5.5 Detection of truths and the fidelity of facial expression

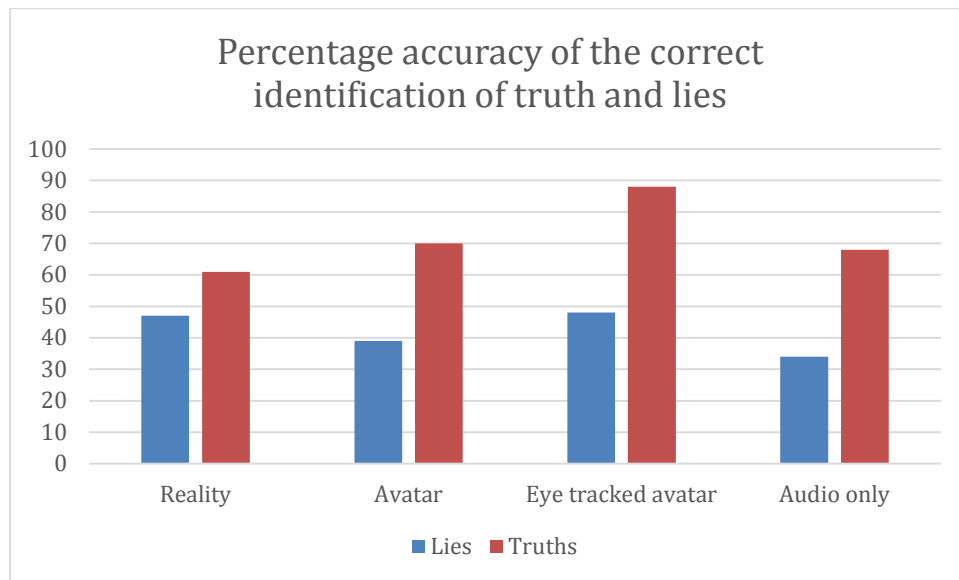
Detecting deception is of great importance across several fields including within judicial, security and law enforcement services. In studies using law enforcement personnel, accuracy rates when detecting truths range from 45-60% (Vrij and Mann, 2001), and media quality has been demonstrated to significantly impact upon detection rates (Horn, 2001). It has been shown that training in deception detection can significantly increase accuracy (Frank and Feeley, 2003) and it has been suggested that virtual environments may be suited to deception detection training (McKenzie et al., 2003). Research has been conducted by psychologists

into the identification of deception using facial expressions (Ekman and Friesen, 1969, Ekman, 2003, Feldman et al., 1979) and VR groups have studied lie tracking using eye-tracked avatars (McDonnell et al., 2012, Steptoe et al., 2010). These previous studies have proved to be both an inspiration and a building block for this research.

2.5.5.1 Accuracy of determining truthfulness in the natural and virtual world

When an individual has a tendency to over trust/believe someone is telling the truth regardless of a statement's veracity, this is referred to as truth bias (Millar and Millar, 1997). During lie detection people tend to over trust others in both the real (CF Jr Bond, 2006) and virtual world (McDonnell et al., 2012, Steptoe et al., 2010). The detection of deception without any training is a difficult task and most people are poor at it. In an extensive meta-analysis Bond and DePaulo (CF Jr Bond, 2006) synthesised the results of 206 studies involving 24,483 receivers. This analysis reported an overall accuracy of 54% yet, interestingly, there was a bias towards believing people were telling the truth, regardless of the truthfulness of their statement. The researcher has found only two studies using virtual avatars that present data comparable with Bond and DePaulo's study (CF Jr Bond, 2006). Both studies used motion tracking to drive head and eye movement. The study that additionally used motion tracking to drive facial expression (McDonnell et al., 2012)) reported a truth bias comparable with that of the study by Bond and DePaulo (CF Jr Bond, 2006). The study that did not use motion tracking to drive facial expression (Steptoe et al., 2010) reported an exaggerated tendency to trust. In Graph 1, the correct identification of truth and lies from Bond and DePaulo (CF Jr Bond, 2006) and Steptoe's studies are shown (Steptoe et al., 2010). Unfortunately, the researcher could not directly contrast these results with the work of McDonnell et al. (McDonnell et al., 2012) because accuracy breakdowns

were not available, however, McDonnell et al. did report a normal truth bias similar to that shown in previous psychological literature. Due to the small number of studies identified and varying experiment designs a firm conclusion cannot be drawn, however, it is interesting and more studies utilising avatars are needed to investigate this.



Graph 1. The percentage accuracy of the correct identification of truth and lies. Reality values taken from Bond and DePaulo's meta-analysis (CF Jr Bond, 2006) and all other values taken from the study conducted by Steptoe et al. (Steptoe et al., 2010).

2.5.5.2 Measurement

Both the above studies measured performance in terms of the accuracy of determining truthfulness, split between true and false statements. To tie this to the impact of quality of facial expression, it would be interesting to quantify what facial expressions each medium could show. To strengthen this link it would be useful to know how hard people had to work

to detect deceit in both mediums. The section below summarises related work in the identification of facial expressions and the methods of measuring cognitive workload.

Facial expressions, it is suggested, are in-built human behaviours which occur in all cultures (Ekman, 1993). Micro expressions are facial expressions that occur very quickly. Humans cannot control them and they happen in a split second, leaking information about feelings (Ekman and Friesen, 1969). Previous studies on micro expressions show that they may allow medical doctors to better understand the needs of patients by identifying emotions such as fear or sadness (Endres and Laidlaw, 2009). With the push towards telehealth, a practitioner's ability to identify patient emotions through media could be important. Micro expression tuition is being researched as a way of helping military and law enforcement personnel to detect deception (Slater, 2009, Porter and Ten Brinke, 2008). Micro expressions may also have applications in the judicial system (Bernstein and Loftus, 2009). A facial coding system has been created to map and identify micro expressions (Ekman and Friesen, 1978) and previous studies have investigated how accurately facial expressions and emotions can be identified using avatars (Dyck et al., 2008).

The purpose of mental workload measurement is to assess the effect of cognitive demands that are associated with task performance. Previous studies have demonstrated that high cognitive workload is positively correlated with greater neuronal activation in the prefrontal cortex (Ayaz et al., 2012), as measured by functional near-infrared spectroscopy (fNIRS). fNIRS is a non-invasive optical neuroimaging method for mapping the functioning of the human brain. It offers better spatial resolution than EEG and more temporal resolution than fMRI (functional magnetic resonance imaging) (Ferrari and Quaresima, 2012). fNIRS is also, arguably, less likely to influence results than fMRI due to its low encumbrance, the

fact that it does not enclose the subject, and its quick set up. These factors make it a promising neuroimaging technique for investigating cognitive workload in virtual environments and, indeed, it has already been used in previous studies (Hu et al., 2013, Seraglia et al., 2011, Cansiz and Tokel, 2012).

2.5.5.3 Approaches to capturing the face and reproducing it in 3D

This section considers the various approaches used for capturing and reproducing facial expression, identity and appearance.

Approaches to capturing facial expressions for authored avatars either track specific parts of the face such as eyes and mouth, or the face as a whole. The former is well suited to driving an avatar face that is visually different to that of the owner. For the latter, a model of facial muscles within the avatar is driven by motion tracking data. Both marker based (Vicon) and markerless (Faceshift) capture approaches are popular. Marker based tracking is more accurate, whereas markerless tracking does not require large numbers of tracking markers stuck to the face. The ability to accurately track the muscle groups of the face is significant as a smile does not only affect the mouth, it also engages muscle groups near the eyes (Porter and Ten Brinke, 2008). Without the muscles of the face being mimicked it might be more difficult for viewers to identify emotions. However, driving the simulation of muscle movements is not the only approach to mimicking facial expression.

Other approaches, such as those used for telepresence (Fuchs et al., 1994, Raskar et al., 1998) derive a fluid 3D reconstructed avatar from live video. These approaches attempt to faithfully reproduce appearance, not just movement. Popular 3D reconstruction approaches

in this area are shape-from-silhouette (Roberts et al., 2013a) and depth-based (Maimone and Fuchs, 2011, Kinect). Both impact on the quality of reproduction in different ways. Shape-from-silhouette cannot model concavities, such as eye sockets (although texture information can simulate this). However, when cameras are satisfactorily calibrated together, the visual-spatial quality increases with the number of cameras. The predominant depth-based approach (Maimone and Fuchs, 2011, Kinect) bounces infrared structured light off the subject. The resolution of the depth-map is thus determined by the resolution of the structure reflecting off the subject rather than that of the camera. Depth-based cameras are also used to drive pre-modelled avatars, including the mimicking of facial expression (Faceshift).

Abstract avatar approaches can hide identity and details of appearance. Such hidden details include perspiration and small facial movements which are not reflected by marker driven models of muscles, although change in skin tone (blushing) has been modelled (Pan et al., 2008). All these details can be captured using 3D reconstruction approaches. Abstract avatars are suited to applications where protecting anonymity is important, thus increasing a person's confidence and perhaps giving them an unprejudiced hearing. In the real world a person's physical appearance can impact upon how honest one perceives them to be (Spellman and Tenney, 2010). It should, therefore, be considered that avatar representation can impact upon trust. 3D reconstructed avatars are suited to giving an accurate portrayal at the expense of anonymity. They could, for example, be well suited to virtual identification parades (Bailenson et al., 2006a). This has the advantage of allowing witnesses to identify criminals without being present and the task could take place in a virtual crime scene, perhaps therefore helping with recollection.

2.5.5.4 Validity of comparing video based 3D reconstruction to video.

The aforementioned meta study of the detection of deception (CF Jr Bond, 2006) includes many studies that have used video to ensure that participants are given consistent data. High quality video captures the vast majority of facial expressions. Facial action coding (Ekman and Friesen, 1978) typically uses video both for repeatability and for frame by frame analysis. 3D reconstructed video provides a fairer comparison with video conferencing, than with an abstract virtual medium, as all modelling is live. A medium can be more easily compared to video if its contents are derived from that video and attempts a faithful reproduction. Firstly, exactly the same input data can be used for 2D and 3D video, albeit only one channel can be used for 2D video. Secondly, any visual difference is an unwanted consequence of the reconstruction process rather than a wanted abstraction. These reconstruction problems may include, but are not limited to, slicing, droop, texture mapping issues and colour matching errors. All of these pose potential challenges when identifying facial expressions. Slicing can cut off relevant detail, droop caused by camera positioning can change visual appearance and can make one emotion appear like another, perhaps making the capture subject look sad (Roberts et al., 2015a). Poor texture mapping impacts on visual quality and inadequate colour matching could cause a reconstruction to look over exposed or too dark.

2.5.6 Summary of the discussions in section 2.5

In virtual reality and gaming, synthetic characters are used as agents (non-player characters) or avatars (representation of the user). These characters need to be able to represent human behaviour to a standard sufficient for their purpose, be it for entertainment (Activision, 2010)

or for a more serious application (Bekele et al., 2013). Two measures of virtual characters are realism and believability. The quality of a character's appearance and behaviour can impact upon the perception of them (believability). The theory of the Uncanny Valley, created by Japanese roboticist, Dr Masahiro Mori, (Bryant, 2005, Groom et al., 2009) suggests that there is a correlation between user perception and realism. Once a character gets close to a realistic appearance, but not quite, it can become somehow sinister, or less believable than a much simpler cartoon version of itself. This evident in movies utilising CGI characters (Gutierrez, 2011), with some filmmakers producing less human characters to avoid the uncanny valley (Brenton et al., 2005, Weschler, 2002). The same approach has been taken in gaming using toon rendering (Brenton et al., 2007). It is suggested that visual appearance and behaviour also need to match so that characters do not cause unease (Brenton et al., 2005, Tromp et al., 1998).

The literature suggests the avatar that is worn by a user will have an impact on their behaviour within the virtual world (Peña et al., 2012) (Fox et al., 2013) and can influence other's opinions of them (Behrend et al., 2012), possibly also temporarily affecting behaviour outside the VE (Rosenberg et al., 2013). This would suggest that users do, indeed, need to be thoughtful when picking an avatar. Another issue highlighted is how an avatar's appearance can influence the perception of a character's empathy towards the user (Dean et al., 2009) and the user's own self-disclosure to the avatar (Bailenson et al., 2006b). In the real world an individual's appearance can impact upon another's trust of him/her (Spellman and Tenney, 2010) which would suggest that more thought should be given to how an abstract avatar's appearance can impact upon trust. For example, if an avatar is to represent a witness in court, to protect their identity, it is important that the avatar does not influence the jury's interpretation of their testimony, or increase truth bias. Some prior studies have

looked into the use of VEs in courtrooms but did not look into the use of virtual characters in any great depth (Bailenson et al., 2006a).

Previous studies have investigated how accurately facial expressions and emotions can be identified when using avatars (Dyck et al., 2008) and psychologists are researching the use of micro expression tuition as a way of helping military and law enforcement personnel detect deception (Matsumoto and Hwang, 2011, Porter and Ten Brinke, 2008). VEs have been shown to be an effective way of teaching cultural NVB (Babu et al., 2007) and CGI has been found to be comparable with video for identifying behaviour in both people (Smith et al., 2012b) and animals. Micro expressions, it is suggested, are in-built human behaviours which occur in all cultures (Ekman, 1993). They cannot be controlled and happen in a split second, leaking information about our feelings. They may allow medical doctors to better understand the needs of patients, for example, expressing fear or sadness (Ekman and Yamey, 2004, Endres and Laidlaw, 2009). Juries could potentially use micro expressions to help them decide whether a person is being truthful (Bernstein and Loftus, 2009). Does the existence of these micro expressions require application developers to rethink the way in which they represent avatars' and agents' NVB within serious games?

In an emergency planning simulation, if a remote member of the personnel displays a micro expression suggesting fear, is it not important to portray this? Such an expression could be a signal of danger, worry or even doubt that a command may not be the correct thing to do. All such expressions are important displays of NVB and all such expressions could be lost depending on the method of data capture and visualisation. If micro expressions are to be used, then experiments should be conducted to test mediums' ability to capture and display them upon virtual characters. It is suggested that one way of measuring a medium's ability

to do this is to compare its ability to display facial muscle movements with that of video (commonly used for FACS coding), therefore using FACS (Ekman and Friesen, 1978) as a measure of character facial movement quality. If a medium can portray muscle movements on a par with video, then that medium can, in turn, display micro expressions (which are muscle movements) assuming that there is a high enough frame rate to capture them.

The visualisation of facial expressions has applications outside that of virtual humans. Other members of the primate family also display NVB through facial expressions. The Chimpanzee, for example, can show several emotions in this manner. Research has already been conducted looking at reproducing Chimpanzee facial expressions using 3D models and a Chimpanzee facial action coding system exists (Parr et al., 2007). This type of research may well help researchers better understand the evolution of facial expressions (Parr and Waller, 2006). If visualisation techniques can be applied to people, they can also be applied to other species to help facilitate the learning of animal behaviour or, in the case of Chimpanzees, it could even allow communication with an animal through a Chimpanzee avatar that displays Chimpanzee behaviour and facial expressions, thus allowing researchers to better understand how Chimpanzees communicate.

If it is an aspiration to enable virtual characters to display human behaviour accurately, then tests need to be conducted to discover how well people can identify behaviour from them, and to explore which more subtle behaviours would be required to identify intent. Gross movements, such as a walking or running, can have a different meaning depending on subtle facial expressions, changes of speed, or hand gestures. Research in these areas would enhance virtual social interactions. The literature suggests reproducing facial expressions on avatars reduces the tendency of observers to over trust (McDonnell et al., 2012, Steptoe et

al., 2010) although there is only a small number of studies in this area and more work would be needed to confirm this. However, no studies exist yet for the medium of video based 3D reconstructed avatars. It would, therefore, be interesting to test this medium alongside the others and see if a truth bias is present.

2.6 Summary

It is an exciting time for virtual reality with the costs of hardware (Oculus, Kinect) and software coming down and making its use financially viable for research, development and home consumers. Educational virtual environments/serious games are a growing market with companies recognising virtual reality's capabilities for training and simulation. Virtual characters are a key area of research within virtual reality often acting as facilitators for learning or communication, as either agents or avatars, but currently representation of non-verbal behaviour is limited. Work is needed to ensure standards are set for the creation and use of virtual characters to enable them to display sufficient behavioural realism to allow for the identification and interpretation of behaviour.

Serious games offer a new approach to tackling sometimes difficult to teach subjects (these subjects can be difficult to teach because of expense (Ieronutti and Chittaro, 2007, Falloon, 2010, Dalgarno and Lee, 2010) or because their tasks are dangerous to approach without prior training (Rizzo and Kim, 2005). Virtual characters can act as guides or storytellers steering users towards learning outcomes (Sims, 2007) or they can represent a user in an immersive virtual environment (Vinayagamoorthy et al., 2004). VR and serious gaming can be used to train medical professionals (Graafland et al., 2012), can virtually teach police officers to conduct interviews (Frank et al., 2002) or can allow remote collaboration on tasks

(Roberts et al., 2015b) in a manner more akin to the real world. In all of these tasks the ability to convey NVB is of benefit.

With regard to animal behaviour, the literature suggests that zoos have an obligation to teach visitors about their exhibits (WAZA, 2005) and need new and innovative ways of doing so. Research has identified that certain animals respond positively when keepers care for them by mimicking the animals' behaviours (Jenvoid, 2008, Case et al., 2015), and some research has been carried out in these areas on teaching primate behaviour (Hay et al., 2000) (Crowell, 1997). Such scenarios illustrate that serious games are a way of teaching material that may otherwise be too expensive (Ieronutti and Chittaro, 2007, Falloon, 2010, Dalgarno and Lee, 2010) or too dangerous to teach (Rizzo and Kim, 2005). However, before moving onto undertaking research on more advanced studies, there is a need to examine whether CGI used in VEs contains sufficient detail for the identification of animal behaviour. By fixing viewpoint and scripting the user experience in a VE, it is possible to compare this experience with that of viewing film, in essence isolating the visual elements of both types of media, allowing a fair comparison.

In virtual reality and gaming synthetic characters are used as agents (non-player characters) or avatars (representation of the user). Two measures of virtual characters are realism and believability. The quality of character appearance and behaviour can impact upon our perception of them (believability). The theory of the Uncanny Valley (Bryant, 2005, Groom et al., 2009) suggests a correlation between user perception and realism. When a character gets a near realistic appearance, but not quite, it can become somehow sinister, or less believable, than a simpler cartoon version of itself. This effect can be seen in movie characters (Gutierrez, 2011) with some filmmakers reducing visual realism to avoid the

uncanny valley (Brenton et al., 2005, Weschler, 2002). This also occurs in games (Brenton et al., 2007). It is suggested that visual appearance and behaviour need to match so that characters do not cause unease (Brenton et al., 2005, Tromp et al., 1998).

It could be suggested that while visual realism has improved enormously in the past ten years, behavioural realism is taking longer to achieve the same level of success. In 2006 Vinayagamoorthy et al. (Vinayagamoorthy et al., 2006) produced a state-of-the-art paper about building expressions into virtual characters. In computer games it is evident that many of the areas discussed are still not being implemented commercially and characters are still somewhat “wooden”, as the authors stated in 2006. This is evident in avatar communication, with older titles initially using text based communication before switching to voiceover IP (Halloran et al., 2003, Halloran et al., 2004). Social presence is impacted upon by communication bandwidth (Peña and Hancock, 2006), so the more communicative cues that are available, the better. The inclusion of NVB, such as facial expressions, could, therefore, enhance social presence. It is known that authored avatars can be driven using depth cameras (Faceshift, 2015) such as the Kinect (Kinect). It is, therefore, not a huge leap to posit that player facial expressions could one day be transmitted in real-time to their avatar in multiplayer games. However, if there is a push towards higher NVB in gaming or VR and, in particular, in visually immersive applications, there is also a potential hurdle in the current display technologies. It is not, presently, possible to capture facial expressions while wearing HMDs, nor can many people afford, or have the room for, a spatially immersive display. Even if this were possible, stereoscopic glasses would occlude some parts of the upper face, thus currently the capture of facial expressions in visually immersive technologies is problematic. One possible solution is video based 3D reconstructed avatars (Roberts et al., 2015a).

The studies which have been discussed show that both avatar users (Peña et al., 2012) (Fox et al., 2013) and viewers (Behrend et al., 2012) are influenced by avatar appearance. Should developers be looking at introducing more NVC from users so as to better represent them? In serious games, in particular, this would be very useful. As previously discussed, a large amount of human communication is through NVB and, in emergency situations, all communication could be important. One area of particular interest is facial expressions. What effects do render styles have on identifying facial expressions? Are they as effective as video? Can users take the same information from authored abstract and 3D reconstructed avatar representations as they can from video?

To summarise, this literature review adds knowledge to the area of serious games, identifying that zoos need a better way of informing visitors about animal behaviour (Davies, 2000) which could also contribute to conservation education which is a main aim of zoos (WAZA, 2005). In addition, the review identifies that zoologists could benefit from a better way of learning about animal behaviour and, in particular how to mimic it (Jenvoid, 2008, Case et al., 2015). The researcher suggest that the creation of animal behaviour training applications could be of use to both the general public, zoology students and zoo staff. It would be of utmost importance to ensure that the virtual animal visual realism was to a standard whereby viewers would be able to identify animal behaviour. Without this they would not be fit for purpose.

This review contributes to the area of virtual characters. The researcher has identified a possible trend across the literature that suggests reproducing facial expressions on avatars reduces the tendency of observers to over trust (McDonnell et al., 2012, Steptoe et al., 2010). It should be noted, however, that due to the small number of studies available, more

work will be needed in this area to prove this, while, in addition, more work should be undertaken to identify whether facial action coding (Ekman and Friesen, 1978) could be used as a way of comparing the quality of user representation across different media, rather than its normal use to code images. There are already studies which have used real people (CF Jr Bond, 2006) and authored avatars (McDonnell et al., 2012, Steptoe et al., 2010) for lie detection, but no studies yet exist in the field of video based 3D reconstructed avatars. It would, therefore, be interesting to test this medium alongside the others and see if a truth bias is present. It is important to identify if truth bias is exaggerated in avatars with lower behavioural realism because this would allow users to potentially pick an avatar with a lower behavioural realism in order to gain a trust advantage. If this was used in conjunction with other advantages that are possible through avatar visual appearance (Peña et al., 2012) (Fox et al., 2013), it could become a transformed social interaction (Bailenson et al., 2004, Bailenson and Beall, 2006). This could have both positive and negative impacts on avatar mediated communication.

3 Chapter 3 - Experiment 1: Gross behaviour identification – a comparison of video and CGI for the identification of Chimpanzee behaviour.

3.1 Introduction

Knowledge of social interactions between animals has traditionally been passed on through standard teaching materials such as lectures, books, documentaries and field work. Serious games may be well suited to teaching animal behaviour as they allow easy control of what is presented, and interaction, both in moving around the environment and in effecting the behaviour of the virtual animals. This study compares how the visual qualities of computer generated imagery (CGI) and video on its own impact on identification and enjoyment. This has been undertaken partly because other work has shown how some human behaviours can be more easily communicated using CGI (Smith et al., 2012b) and also to allow the impact of visual qualities to be better understood before studying the interaction.

The identification of the social behaviour of humans and animals from visual mediums has a number of potential uses, such as training through to psychology. Past studies have compared film with interactive virtual environments (Morgan et al., 2002), however, the impact of visual differences in the media on behaviour identification has been studied less. It is known that more detail is often represented in film, whereas unnecessary detail can be left out in CGI. Behaviour, such as the eye gaze of a child with Attention Deficit Hyperactivity Disorder (ADHD), has been shown with great clarity using simplified virtual representations (Adams et al., 2009); furthermore, scenes can be tailored to fit an application rather than exactly mimic reality. Rizzo and Kim state that, in some instances, it is not required that VR exactly copy reality, due to constraints such as time and safety (Rizzo and

Kim, 2005). This report uses virtual Chimpanzees to depict various social behaviours. This allows the user to view a simplified representation of events without being distracted by secondary actions or unnecessary detail within the scene. The aim of this experiment is to measure the relative performance and experience of identifying animal behaviours from CGI and video.

3.2 Method

A Chimpanzee keeper was motion captured mimicking Chimpanzee behaviour and the animations were applied to virtual Chimpanzees. Film was then captured of captive Chimpanzees (real film). Behavioural animations were matched to the real film for the virtual Chimpanzees and footage captured of this (CGI film). Both films were then condensed to less than three minutes in duration. These films were shown to n=20 participants who were asked to identify the Chimpanzee behaviour. Participants were split into two groups of ten. One group was tested via a standard laptop, the other group via a large projected display. In each group, half the participants watched the CGI film first and the other half watched the real film first. The experiment used a repeated measure within subjects' design for the primary overarching task of behaviour identification and between subjects' design for the influence of display. The duration of the experiment was approximately twenty minutes.

3.2.1 Thesis and Hypothesis

The thesis of this experiment is that CGI and virtual animals can be created to represent animal behaviours, mimicking those seen in real life and captured on film. The interactive

aspects of virtual environments are put aside in order to examine how the visual qualities of CGI and film on its own impact on users' ability to identify behaviour and their enjoyment while doing so. The following hypotheses were specifically tested:

H1: There will be no significant difference in the identification of simple behaviour across the two visual mediums.

H2: The participant will perceive the visual fidelity of the CGI film to be high enough for basic behaviour identification.

H3 The participants will perceive the real film to be more enjoyable than the CGI film.

3.2.2 Scope

This study examines how participant experience is affected by the use of different visual media, the first trial being a film that is a faithful representation, the second trial a film captured from a virtual environment (CGI film), that is, an abstract representation. During the experiment, behaviour identification and the participants' impression of visual fidelity and enjoyment were measured. In addition, participants could comment on each medium. Thus, the experiment allowed for the measurement of the visual suitability of a virtual environment for the identification of animal behaviour.

3.2.3 Variables

The dependent and independent variables are discussed below.

Independent Variables

The conditions of the experiment used visual media and display technology. The visual medium variables were CGI film and real film. The variables for the display were a laptop (low immersion) and a large projected display (high immersion).

	Standard Laptop	OCTAVE (single screen)
Real Video	RVI	RVo
CGI Video	CGII	CGIo

Table 4. Independent Variables

Dependent Variables

Three dependent variables were collected. Firstly, there was behaviour identification, where a simple scoring system was used. A correct identification scored 1 and an incorrect identification scored 0. A maximum accuracy score of 8 could be achieved if all were correctly identified. In each medium, seven of the eight behaviours were present and one was not. Secondly, visual fidelity and thirdly enjoyment data were collected on a bespoke questionnaire using a seven point Likert scale with a balanced number of questions and counter questions for each.

3.2.4 Procedure

Each participant read about several Chimpanzees' behaviours using written descriptions taken from an ethogram (catalogue) of Chimpanzee behaviour. These behaviours were

selected by a group of primate experts from The University of Salford's School of Environment and Life Sciences. Each participant was shown two short films, one was footage of real captive Chimpanzees, the other was of virtual Chimpanzees. Both films were approximately two and a half minutes long and illustrated various Chimpanzee behaviours. Participants then had a second chance to review the written behavioural descriptions before again watching the two films. They were permitted to make notes as they watched the films. On completion they were given a questionnaire that allowed for the collection of objective measures of task performance and subjective measures of experience. Participants were asked to record behaviour identification only if a behaviour had been viewed at least once.

3.2.5 Subjects

N=20 participants from the University of Salford staff and students were recruited to take part in the study; 14 were male and 6 female. It should be noted that it would have been better to recruit an equal number of male and female participants. The mode age range was 21-30 years (45%). Participants were not paid for taking part. The majority of participants either never, or only occasionally, played games (80%). Participants were free to leave the study at any point with no explanation needed. It was explained that the results of the experiment were anonymous and that the participants were not in competition with each other. Only the researcher and his research partners would see the results. Participants were not rushed at any point and were free to ask the researcher any questions.

3.2.6 Ethics

Ethical approval was gained for the study from the University's Ethics Board. The approval number was CST 12/51. The experiment's participants were given an information sheet and a consent form. The information sheet explained to the participants the general theme of the study and what they might expect to see. The consent form informed the participants of the anonymity of data, of their right to leave at any time during the experiment, and of their right to have their data completely removed.

3.2.7 Apparatus

The techniques used for the creation of the CGI film and the real film are described below. In addition, there is a brief description of the display technologies used.

Experimental Virtual Environment and Video creation process

In order to test the value of using VE technologies as an aid in teaching animal behaviour and in contributing to conservation education, an experimental VE, firstly, had to be built. This could then act as the test bed for the behaviours and allow user performance in the environment to be measured against performance in another medium. An experiment was used to test the behaviour identification affordance of the VE medium (CGI film) and to compare it with a standard film medium, thus allowing the basic performance of the VE to be measured independently of interaction and then contrast the user experience with that of the standard film.

Modelling – Modelling for the project was undertaken in Autodesk 3DS Max (Autodesk, 2015a). The environment and the virtual Chimpanzees were produced using low polygon modelling techniques (4694 polygons per Chimpanzee) in an effort to keep up performance while maintaining adequate visual realism. The idea was to keep the polygon count low enough so as to enable the model to run in real-time at a sufficient frame rate. If the frame rate is allowed to drop to low then the environment starts to stagger and an effect called lag can be seen, thus diminishing the user experience. Once the characters were created they were given a bones' system which was used to deform the skin mesh when the characters were animated. They were then exported into Autodesk motion builder in order to have some motion capture data assigned to them. This motion capture data would bring them to life.



Fig. 13. Virtual Character

Texturing – All textures for the project were created using Adobe Photoshop (Adobe, 2015). The texture sizes were kept to a minimum that allowed for optimum real-time performance whilst maintaining an acceptable level of detail. All textures are square and saved in a dds file format. The use of smaller more optimal textures improves loading times and performance.



Fig. 14. Virtual Chimpanzee

Motion Capture – Chimpanzees are very dangerous animals. For safety, and so as not to upset the animals, it was decided that a zookeeper would be motion tracked acting out the Chimpanzee behaviours. As keepers have first-hand experience of the animals, and are familiar with their behaviours and movements, this was seen to be the optimal way to capture their motions. The MOCAP for the project was conducted using a Qualisys Motion Capture Facility (Qualisys, 2015) using a VICON suit (Vicon, 2015). The data was then cleaned and imported into Autodesk Motion Builder where it was placed upon a character. Once ready, this character could then be exported back into 3Ds Max (Autodesk, 2015a) and then into Quest3D (ACT3D, 2015) for real-time output.



Fig. 15. Virtual Chimpanzee creation process from the motion capture (left) then via the untextured to the textured geometry

Animation – Animation for the project was created using motion capture data and key framing techniques. The three pieces of software used for this were Qualisys (Qualisys, 2015), Autodesk 3Ds Max (Autodesk, 2015a) and Motion Builder (Autodesk, 2016).

Rendering Engine – When looking at the best way to create the simulation, it was apparent that time would be a limiting factor. There would only be one person developing the project (Model creation, motion capture, animation and programming). This meant that an engine had to be chosen that was not only powerful and flexible but also could provide swift development. Quest3D (ACT3D, 2015) is a real-time rendering engine that allows for visual programming. It also has an application programming interface (API) that allows the user to hand code any additional channels that are needed. It is a powerful and flexible tool and was the ideal choice for the project.



Fig. 16. Virtual Enclosure

CGI film production

The CGI film was created by screen capturing footage from the experimental virtual environment. This was edited into a short sequence with no audio.



Fig. 17. An image showing the virtual Chimpanzee enclosure

Real film

Film footage was captured showing Chimpanzee behaviour. This was edited into a short sequence with no audio. It should be noted that it was a time consuming process waiting for the Chimpanzees to display particular behaviours and there was no guarantee that this would happen which, in itself, could add weight to the argument for some kind of virtual training application.



Fig. 18. Approach behaviour.



Fig. 19. Eating behaviour.



Fig. 20. Quadrapedal Walk

Hardware/Display technology

Display Variable 1 used a gaming laptop. Display Variable 2 used a large projected screen within the OCTAVE (Roberts et al., 2013a) allowing for a life sized, first person experience.

3.2.8 Questionnaire design

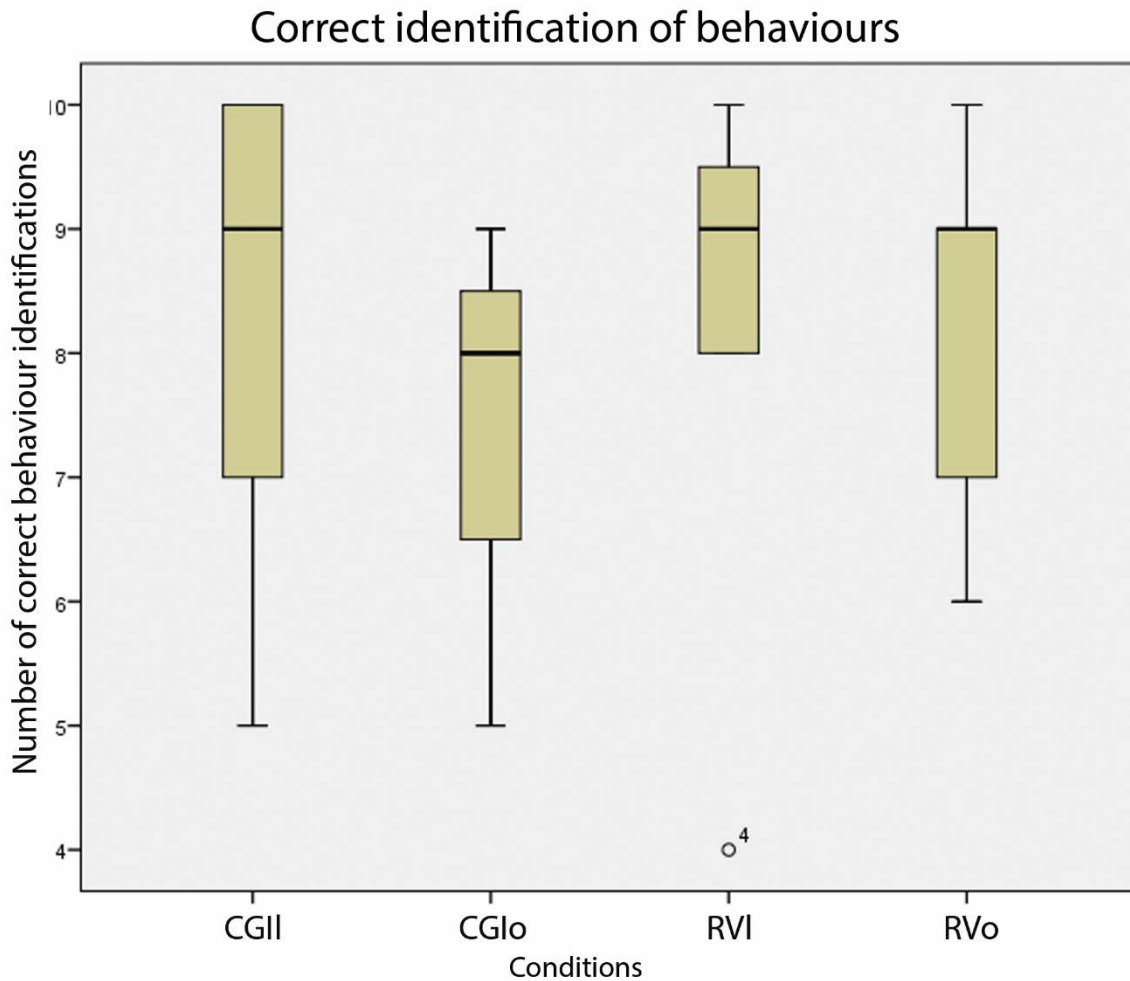
A questionnaire was designed and split into two sections; the first section contained a brief personality questionnaire and the second section posed questions related to the experiment. The personality questionnaire determined a participant's suitability for the experiment. The experiment section contained a behavioural identification section. Participants could tick off the behaviours that they had identified in each of the films. In addition, an experience section consisting of twenty questions (including counter questions) measured enjoyment, presense and visual fidelity. A seven point Likert scale measured participants' opinions.

3.3 Results

3.3.1 Behavioural identification performance

A McNemars test was used to compare performance across the media. There was a strong trend towards better performance in CGI_{II} (83.8% correct) when measured against CGI_O (75% correct). This was significant ($P=0.039$). Participants performed better in RVI (83.8% correct) than in RVo (82.5% correct). This was not significant ($P=0.9999$). Therefore, only the CGI medium performed better in low immersion. Participants performed equally in RVI (83.8% correct) and in CGI_{II} (83.8% correct). This was not significant ($P=0.9999$). There was a trend towards better performance in RVo (82.5% correct) when measured against CGI_O (75% correct). This was not significant ($P=0.143$). Overall the real video (83.1% correct) performed better than CGI (79.4% correct) but not significantly ($p=0.377$).

There were differences in the qualitative data collected suggesting that the real film was busy and unpredictable, with comments such as “The film required attention to be focused everywhere to witness behaviours, however the VE (CGI Film) could focus on a particular behaviour.” The qualitative data also suggested that the film contained more subtle movements and intent “The key issue is granularity. Gross actions were identifiable in both, but nuance and purpose only really came through in the film. For approach you could see contact between two individuals in the film but VR (CGI Film) did not establish this and the animals could have just wandered closer.”



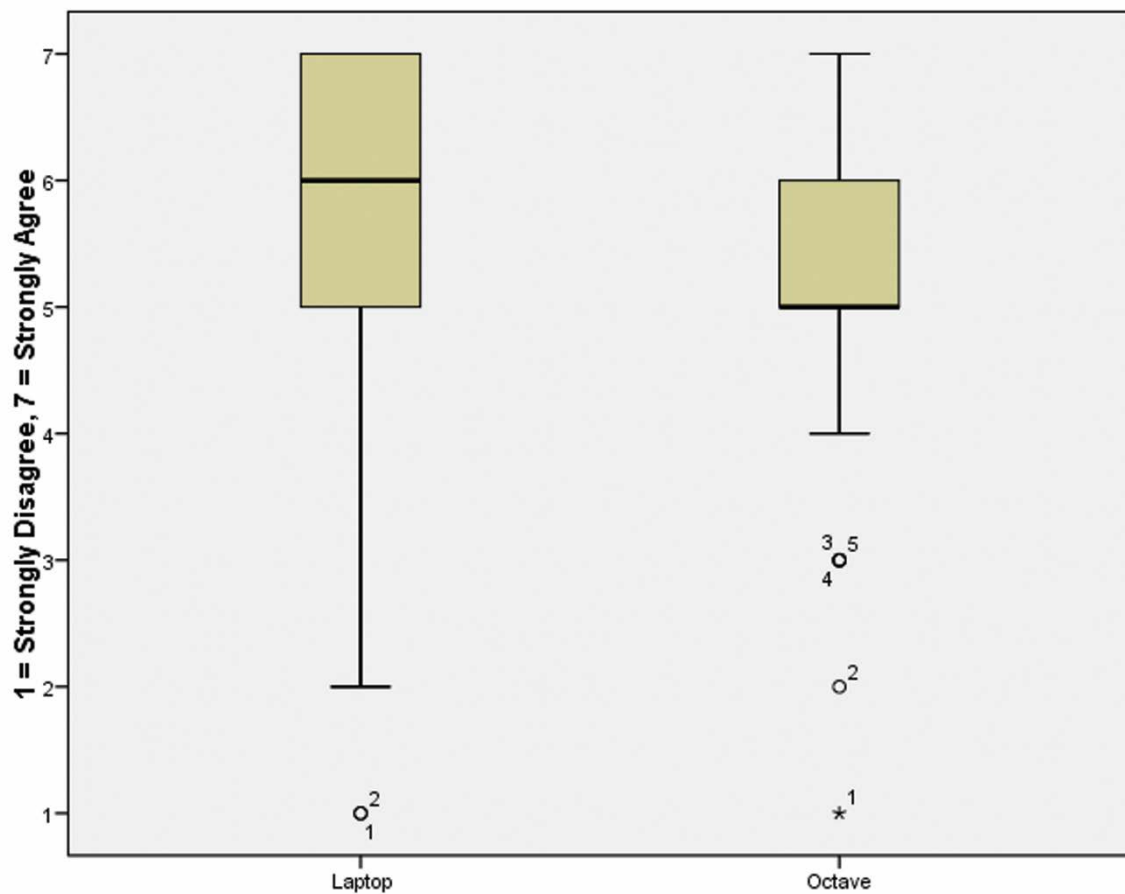
Graph 2. Showing behaviour identification.

3.3.2 Visual fidelity

A Mann-Whitney U Test was conducted to compare participants' impression of the visual fidelity of the mediums. The Mann-Whitney U Test showed a significant difference in participants' impression of visual fidelity levels; CGII results (Md = 6.0, n = 50) and CGIo results (Md = 5.0, n = 50), $U = 861.5$, $z = -2.764$, $p = 0.006$, $r = -0.2764$, with a trend towards better performance in CGII. This suggests that low immersion offered participants a better impression of visual fidelity. Overall 80% of the participants viewing the films on the laptop or on the Octave believed that the visual fidelity was high enough to identify behaviour.

However, the qualitative data suggested that there was room for improvement with participant comments such as “Movement of animals in VR (CGI film) was strange at some points.” Another participant stated “Also the VE (CGI film) has the ape’s movements not so smooth and this was a bit distracting.” This suggests that the movement was detracting from the user experience.

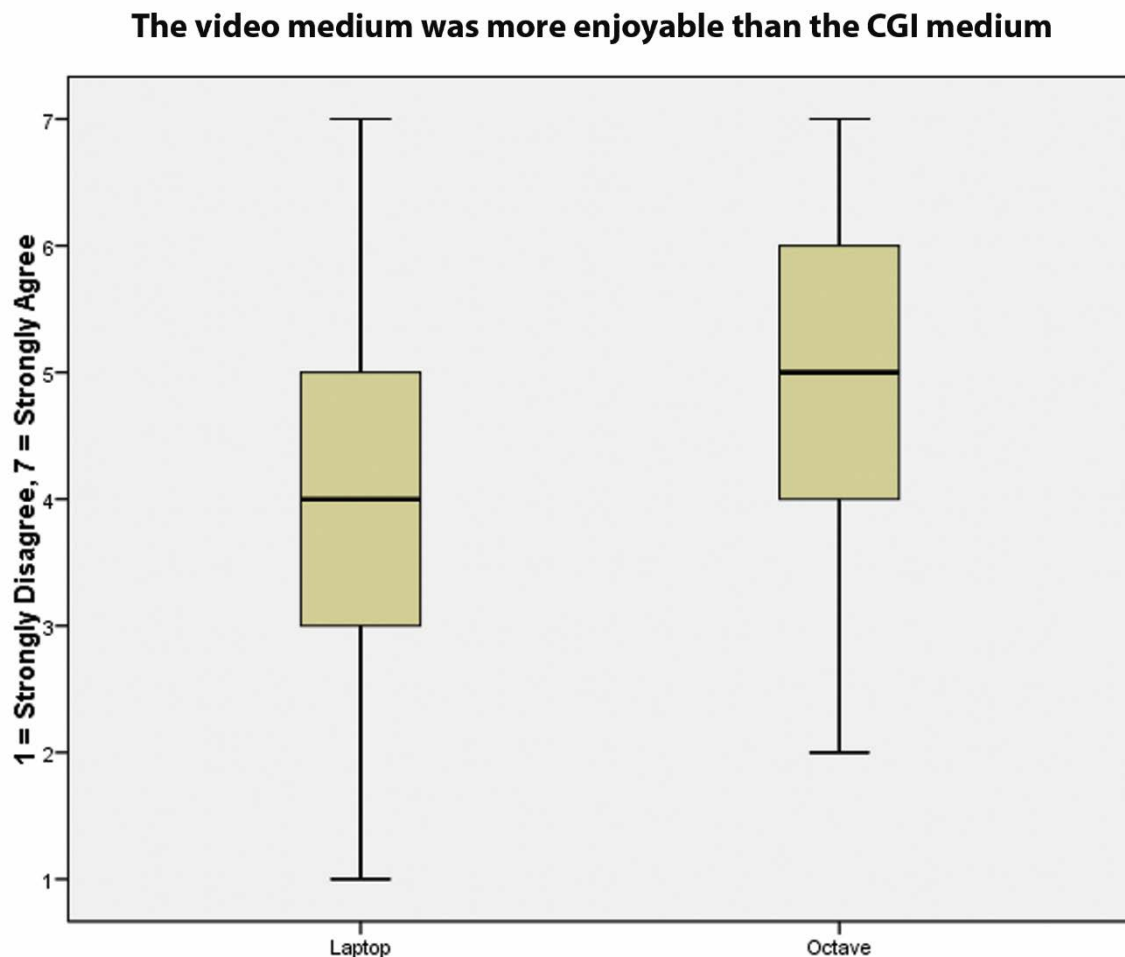
CGI medium visual fidelity is good enough to identify animal behaviour



Graph 3. Whether visual fidelity of CGI is high enough for behaviour identification.

3.3.3 Enjoyment

A Mann-Whitney U Test was conducted for the conditions: ‘the film was more enjoyable on the laptop’ and ‘the film was more enjoyable on the Octave’. The Mann-Whitney U Test showed a significant difference in participants’ enjoyment levels; RVI results (Md = 4.0, n = 40) and RVo results (Md = 5.0, n = 40), $U = 533.5$, $z = -2.611$, $p = 0.009$, $r = -0.29$, with a trend towards higher enjoyment of the RVo. Interestingly, the real film condition was preferred on the larger Octave screen, while the CGI conditions’ performance improved and was equal to the film on the smaller laptop display.



Graph 4. Whether the video was more enjoyable than the CGI video.

3.4 Discussion

The research question set out at the start of the thesis was ‘How do the visual qualities of computer generated imagery (CGI) and film impact on identification and enjoyment?’ The results suggest there is no significant difference between CGI and film for the identification of simple animal behaviour, which was in line with a similar study examining identification of children’s classroom behaviour (Smith et al., 2012a) which was published as this study concluded. This suggests that the CGI film contained enough relevant realism for gross behaviour identification which implies that additional detailing may be a waste of resources and, potentially, a distraction. Further work would be needed in this area to examine this more closely. A possible way of improving this study would be to conduct further experimentation using various levels of detail to try and establish minimum requirements for gross behaviour recognition, allowing a baseline or standard to be established.

Looking more closely at the results for behaviour identification, participants performed better in the low immersion condition (laptop) than in the high immersion condition (Octave); a significant difference was shown in the CGI medium with low immersion outperforming high immersion. This could be attributed to rendering artefacts (aliasing) and pixelation becoming more prominent as screen size increased, while media resolution remained the same; pixelation was also present in the video medium as the screen size increased. However, in the video medium this did not result in a significant difference in performance during behaviour identification.

In qualitative data collected it was noted that the CGI film did not contain sufficient detail to illustrate subtle movements and intent. This is a problem as these subtle behaviours help the viewer interpret meaning behind larger motions. It should be noted, however, that

performance could have been affected by the quality of acting performance given by the zoo keeper which was used to capture the Chimpanzee behaviours, and/or the fidelity at which the motions were captured (number of markers, etc.). Today these subtleties could be added with extra effort and resources such as a trained actors, facial capture, higher resolution motion capture and key framing to add small movements. However, at the time of conducting this study, software tools for facial capture and its real time deployment were not available. Now, it would also be worth considering other mediums for this purpose such as 3D reconstructed video to directly capture the Chimpanzees' behaviour; however, finding an appropriate capture space and blending animation motion sets would be problematic.

Identifying levels of relevant realism for specific tasks has been undertaken in other fields. Biology books utilise media such as line drawings of anatomy or simplified images showing areas demarked by colour in addition to photography. Each type of image is used to help the learning process as the way they are presented makes it easier to identify certain features (Dwyer JR, 1968). This does not occur with CGI. Rendering may occur in a wireframe when when it is desirable to show that not many triangles have been used in a mesh, or to show the topology flow, but there is no guide that tells users which rendering style may be best for a particular style of learning, nor is it known how much detail is needed for optimum learning. In virtual environments unique opportunities are presented to tailor scenes for learning. Details can be removed in order to make scenes uncluttered, or camera paths can be designed that focus the user's attention on certain attributes, whether through model layout or camera depth of field. In effect, it is possible to play with variables and try to find the appropriate level of 'relevant realism'.

The use of CGI allows the user experience to be controlled. In serious games, behaviours and events can be triggered when required, allowing for behaviour to be seen on demand. In the real world, animals choose when they want to do something and quite often take their time in doing so. It is possible to spend hours making observations and never see the desired behaviour. Therefore, serious games could supplement observation training. Past studies suggest that serious games could be used to update, and add to, teaching material (Pan et al., 2006a). However, they need to be fit for purpose and not created for the sake of using technology.

People visit zoos as a form of entertainment (Ross and Gillespie, 2009), so it is important that they enjoy the learning process which can take place in zoos. Participants' impression of enjoyment was split according to which platform the media was viewed on, with the film preferred on the higher immersion projected screen and the CGI film performing equally well as the film on the lower immersion laptop screen. This could, potentially, be due to the rendering artefacts in the CGI film becoming more prominent when viewed at a larger scale or that the increased scale exposed missing detail in the CGI film. A truer test may have been to scale the resolution, as the display scaled, but due to hardware constraints at the time this was not possible. There were also positive comments such as "I think the virtual environment would be a great teaching aid and as children/young people are so exposed to technology it would be easy to include with minimal disruption whilst allowing for information to be passed on in a fun and engaging manner". This would suggest that, with improvement to visual quality for larger displays, the CGI film would be fit for purpose. Overall, the results do suggest that CGI could potentially be used as an alternative to film as a teaching aid to learn about animal behaviours and also within zoos for conservation education, which is one of the main aims of zoos globally (WAZA, 2005).

3.5 Summary

Currently, human and animal behaviour is learnt through observation and standard teaching material. This study suggests that serious games could be a new and innovative way of teaching such behaviour. However, prior to exploring free viewpoint serious games, it was deemed necessary to investigate the impact of visual differences in the media from a fixed viewpoint, thus checking whether any advantage or disadvantage can be identified during free viewpoint studies which would not be due to visual differences in media. The findings suggest that when a viewpoint is fixed people can identify gross animal behaviour equally well in both mediums and previous studies have also shown this for human behaviour (Smith et al., 2012a). In addition, the findings suggest a trend towards better performance in low immersion. However, imagery being scaled up with no increase in resolution could potentially have impacted upon this finding. This work validates the study of free viewpoint interaction and whether it is a benefit or distraction when learning about animal behaviour. In conjunction with this, it would be interesting to study how gradual increases in the levels of immersion impact on learning and enjoyment and whether there is an optimal immersion level for this. Below is presented a summary of the main advantages and disadvantages identified in this experiment.

3.5.1 Advantages of CGI visuals for portraying animal behaviour

1. CGI allows for the level of detail to be managed, thus stopping users from being distracted.
2. CGI allows behaviours to be split into levels. Basic movements could be taught first and then subtle behaviours could be added in.

3. CGI allows for the tailoring of scenarios. Scenarios can be designed to last for a set duration and to contain all the necessary information.
4. There was no significant difference between the CGI film and real film when learning about basic animal behaviours. (This suggests adequate relevant realism.)

3.5.2 Disadvantages of CGI visuals for portraying animal behaviour

1. CGI lacks subtlety of movement and intent. Such subtleties could be added, to a certain extent, with extra resources. However, some users found additional detail distracting.
2. The CGI environment lacks visual richness. Again, this could be added in with extra time and resources.

3.5.3 Contributions of this experiment

The contributions made by this experiment are listed below.

1. The identification of a possible trend across the literature that suggests zoos need a better way of informing visitors about animal behaviour, thus contributing to conservation education.
2. The identification of a possible gap in the literature that suggests zoologists would benefit from a better way of learning about animal behaviour.
3. Demonstrating that CGI demonstrating low level NVB is sufficient for the identification of gross NVB on a par with video.
4. The identification of a possible issue with the identification of behavioural intent due to lack of subtle NVB.

5. Methodological contribution – motion tracking a zoo keeper who mimicked primate behaviour.

4 Chapter 4 - Experiment 2: Lie detection – a comparison of video and 3D video for the detection of deception

4.1 Introduction

This experiment compared the detection of deceit across video conferencing and a 3D video based computer graphic medium. The purpose was to determine if the process of 3D reconstruction influenced trust by reducing the details of facial expression. The inspiration for this experiment came from a small number of previous studies in the natural (CF Jr Bond, 2006) and virtual world (Steptoe et al., 2010, McDonnell et al., 2012) that suggest a stronger tendency to over trust a person when their facial expression is hidden. However, due to the small number of studies that have been undertaken, more work would be needed in this area before firm conclusions could be made. A virtual avatar that copies head and eye movement but not face movements, could be argued as being akin to a person wearing a mask. Thus the opening research question is: Would a 3D medium that removed this mask, result in a truth bias akin to video and thus the real world? Two confederates (senders) each gave a set of accounts, of which half were true. These were captured and transmitted simultaneously using 2D and full 3D video based communication mediums. Recordings of these sessions were later examined by a set of participants. N=21 participants (receivers) who were asked to determine which accounts were true. Measures were taken of: accuracy at detecting truth and deceit, tendency to over trust and the effort undertaken in determining truthfulness. Data were captured using a questionnaire and functional near infrared spectroscopy. The results showed that, despite an obvious visual reduction in quality, the receivers performed and worked to a similar degree in both mediums. The findings of this experiment are of interest to those developing 3D telepresence technologies and virtual humans, and to those concerned with the trustworthiness of a medium.

Previous studies in virtual reality have discussed avatars that hide true identity and appearance (Bailenson and Beall, 2006). Such avatars could have an application in protecting anonymity and reducing prejudice. However, identity and appearance are relevant in many situations where truthfulness must be determined. Furthermore, spatial aspects of non-verbal behaviour, such as mutual eye gaze, are relevant to the detection of deceit (Ekman, 1985), thus giving a reason to use 3D other than to hide identity. The researcher wanted to know if a 3D avatar, designed to faithfully portray identity and appearance, would have sufficient quality of facial reproduction to avoid an exaggeration of the truth bias. Any method for capturing and reproducing a person in 3D induces errors. Previous truth detection studies have limited these errors by using camera tracking to accurately detect markers on the face, and doing much of the modelling off line. Approaches to faithfully reproducing appearance and identity in real time, especially when not impacting on the naturalness of behaviour, can ill afford such luxuries.

In essence, representation through a pre-modelled avatar, brought to life by motion tracking, could be argued as being akin to wearing a mask. The overarching research question is: would a 3D medium, that removed this mask, result in a truth bias akin to video and thus the real world? In order to faithfully communicate identity, appearance and attention, telepresence researchers at The University of Salford are developing a telepresence approach that uses shape-from-silhouette to reconstruct users from live video allowing people to walk around 3D reconstructed versions of each other. The group have experience of how this approach impacts on facial reproduction, and it is known that it offers a direct comparison to video. Thus a derivative of the research question is: Do the shortcomings of the method for reconstruction contribute to a distortion in truth bias?

When describing the technology approach used, the terms virtuality avatar and virtuality telepresence are used to distinguish from conventional authored avatars and the immersive collaborative virtual environments that use them. The term virtuality is used as these avatars reflect the current real appearance of their owners, in addition to simplifying their movements. A conventional avatar is authored offline and later driven by the tracked motion of its owner. As discrete tracking points drive the simulation of movement, some compare this to a puppet following strings.

Throughout this work the term sender will be used to represent the person communicating (truth or lies) and the term receiver for the person detecting the truthfulness of the communication. These terms are commonly used in the psychology literature on lie detection.

In order to compare the work to 2D video, a viewpoint was matched when recording from the 3D video stream. This allowed time stamped films to be created from each medium, to ensure that participants were given equal representations of non-verbal behaviour during the lie detection, allowing a fair comparison. The introduction of a free viewpoint at this stage would not have allowed this.

4.2 Research Question

The research question is:

How do different avatar representations of the same motion data set impact upon trust?

4.3 Hypotheses

In an experiment a sender gave true and false accounts, each sent through both an avatar reconstructed from multiple videos, and through conventional video conferencing. Receivers then watched video footage of both. It was hypothesised that receivers would perform better using the footage of the video conferencing. The rationale is that the detail of the facial expression would be lost or distorted through the reconstruction process. This would make facial expressions harder to identify and, thus, make it more difficult to detect truthfulness.

Hypotheses:

When receivers observe footage of a 3D reconstructed avatar rather than video conferencing:

H1: The accuracy of determining truthfulness will be significantly less.

H2: Cognitive workload will be significantly higher when determining truthfulness.

While it is expected that the reconstructed avatar will be inferior to video conferencing, the experiment tests whether there will be sufficient visual quality to convey truthfulness.

4.4 Method

This experiment investigates whether avatars, reconstructed in real time using shape-from-silhouette, overcome the reported problem of people trusting CGI avatars too much. To test this, the gap between 3D reconstructed video and video conferencing in portraying deceit is

measured. Two senders each gave a set of accounts of which half were true. These were captured and transmitted simultaneously across two mediums. Recordings of these sessions were later examined by sets of participants. As discussed in section 1.4.3 Detection of Deception, n=21 receivers were asked to determine which accounts were true.

4.4.1 Scope

The scope of this experiment was to test if the impoverished facial representation of 3D reconstructed avatars, results in those using the medium to determine truthfulness, over trusting them. The reason that this might occur is that a lack of facial expressions has been shown to have this effect when using pre-modelled fixed expression authored abstract avatars or voice alone. To isolate the impact of visual reproduction from that of support for spatial qualities, the viewpoint has been fixed for the 3D video capture process, close to that seen through the 2D videoconferencing condition; thus not measuring the benefits of support for mutual gaze and interpersonal distance, instead leaving this for a later study. The study discusses the results together with a small number of previous studies to test what the researcher believes a survey of the literature suggests: that a lack of facial expression possibly increases truth bias. However, this was not tested as rigorously as the researcher would have liked. The reason being that the experiment did not implement a third condition to directly test an authored abstract avatar. This was because, in order to record the facial tracking information needed, cameras would have had to be positioned closer to the face. These cameras would have occluded the capture subject from the 3D reconstruction cameras in the capture space. There are different methods available to capture and reproduce the face's impact on quality in different ways. It is likely that another approach, for example

depth-based reconstruction, might produce different results. The results from this study may only be valid for shape-from-silhouette.

4.4.2 Session capture

In this section the set-up of the hardware configurations and session capture process is explained. Firstly, the environments that the sender and questioner occupied are described, in particular, how the information was captured and displayed in both, alongside a brief overview of the technology that linked these spaces through display and capture. Secondly, the telecommunication systems are described.

The sender environment

Each sender was captured within the OCTAVE display system (O'Hare, 2015) designed for telepresence research. The system uses ten HD cameras positioned at the top of the display walls. Figure 21 shows the capture space with someone being captured within it. Figure 22 diagrammatically illustrates the camera positions in the context of the walls and the capture subject, in addition it depicts the camera cones, from the silhouette of the capture subject from each camera image (Roberts et al., 2013b). Only one screen of the immersive display used displayed the videoconference session; the others provided combined lighting and Chroma keying. The display used at the sender's end of the videoconference was a simulation grade active surface. Each of the eight display surfaces are 4x2m and use Christie S+3K Stereo DLP. They were run in mono with a resolution of 1400x1050 at 102Hz. During capture, defuse and ambient light from the OCTAVE's fourteen projectors (6 ceiling and 8

wall) and defuse spotlights were combined to achieve clear lighting and contrast of the capture subject's face. The room's strip lighting was also turned on.

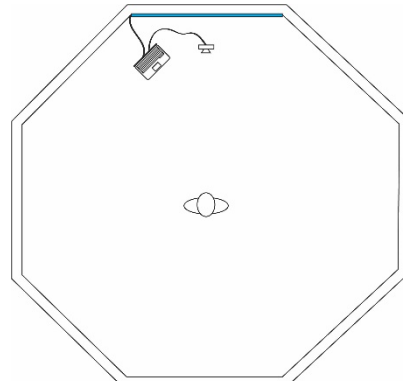


Fig. 21. The sender environment. The sender stands in the centre of the capture suite. A Macbook Pro connected to a webcam captures their image and transmits it through Skype. Simultaneously, multiple cameras (see Fig. 22) are recording and streaming images of the sender for 3D reconstruction.

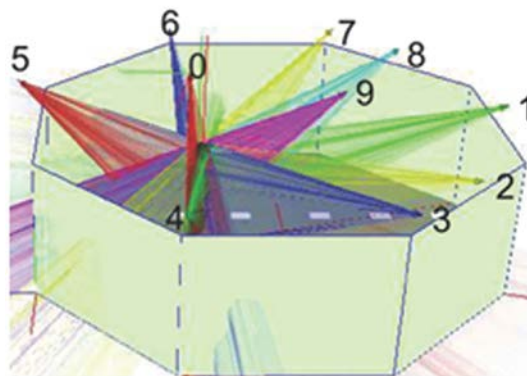


Fig. 22. Camera placement with respect to the sender position and display walls. Silhouette cones joins each camera to the silhouette edges of the sender as seen by that camera. Copyright IEEE (Roberts et al., 2013b).

Telecommunication equipment and set up

The video conferencing and the 3D reconstruction system displayed the head and shoulders of the sender on identically large wide HD screens. Both systems captured the sender simultaneously. The audio from the video conferencing was used for both. There was no discernible difference in the latency of the two systems. Both gave good lip-sync.

The video conferencing software was Skype which, at the time of the experiment, was the prevalent commodity system, widely used at work and home. The version of Skype used was 6.7.0.373. The network was a wireless 1Gb private network under light load. The Skype was run at both ends on Mac Mini computers. These were connected via a 1Gb wireless network under low utilisation.

The 3D reconstruction system created a real-time avatar, utilising both geometry and texture, using shape-from-silhouette. The reconstruction method is outlined in (Duckworth and Roberts, 2014). The capture and pre-processing method is covered in (Moore et al., 2010) while the method of creating form is based on the EPVH algorithm (Franco and Boyer, 2009). The experiment utilised parallelisation to produce real-time frame rates, from 10 HD cameras, on one computer, as described in the study of (Duckworth and Roberts, 2012). The reconstructed avatar was formed, textured and viewed on it. Pairs of video streams were taken by five DELL dual core PCs. These ten streams were then sent to a MacBook Pro in the questioner's room. This laptop undertook image pre-processing, including generating silhouettes, reconstruction and rendering. It achieved this at between 20 to 30 fps dependent on the segmentation complexity of shape-from-silhouette from each camera.

Procedure

Two senders were asked by a questioner to give true and false accounts to four questions. Each sender was asked to walk inside the capture system, stand close to an X on the floor, and face a display wall. Displayed on this wall was a questioner seen in an office through a videoconferencing session. The sender and questioner exchanged pleasantries. The questioner explained to each sender that they would be asked four questions, to which the sender should provide a true and false account. The sender was asked to make each account around one minute long, to randomise the order in which they told truth and lies, and not to tell the questioner which account was true until after all were given. The questioner then asked each question in turn, waiting between each response allowing two accounts to be given. The sender gave these accounts accordingly and, at the end, said which were true. The questions asked were:

Q1. “In under a minute each could you describe two accounts of your route to work today for me please?”

Q2. “In under a minute each could you describe two accounts of a sporting achievement from your past please?”

Q3. “In under a minute each could you describe two accounts of a comical story from your past please?”

One of the following two questions was specifically addressed to each sender in an attempt to add pressure:

Q4. “I have been sending you emails marked urgent and you have not yet responded, could you give me two accounts as to why?”

Q5. “You are in charge of the commercial partnership with body A. Could you give me two accounts as to how this relationship is going?”

Each account was transmitted simultaneously through the two mediums of 3D reconstructed video and video conferencing to two wide screen televisions in the questioner’s room. Video from cameras facing these two screens was used to create 24 short clips.

Subjects

Two senders were recruited. Each worked within the University in roles of commercial engagement. They were chosen as both had excellent communication skills. This was done in an effort to not bias the experiment through using a sender not competent at public speaking or an actor expert in deception through acting. The senders knew the author. One of them knew the questioner.

4.4.3 Detection of Deception

This section outlines the design and implementation of the study to determine the ability of each medium to portray truthfulness.

Procedure:

N=21 receivers watched the series of 24 videos, choosing which were true and false accounts. Each receiver entered a computer lab and sat at a desktop computer. A pad was placed on

their forehead to measure cognitive load. A band holding it was tightened around their head. Wires were attached. A head scarf was wrapped over this equipment to shield it from light and perhaps make the participant less self-conscious. They were asked to follow the instructions on the screen. These told them to watch a series of videos, selecting after each either a button for true or false, based on their overall judgment of the account. They were also asked to press a button each time they thought they detected a lie; however, this data was not used as many reported later to have delayed their decision.

Subjects:

The receivers were recruited from University staff and students: 19 were male and 2 were female. The number of male and female subjects was not matched as they were not interacting with the senders. All were healthy and had the mean age of 37.38 with standard deviation of ± 9.73 years. Recruitment was undertaken via email. Most worked or studied at the University. 17 were staff and 4 were students. The experiment lasted approximately 20 minutes. The participants were not paid for taking part.

Receiver environment:

Each receiver sat in the same computer lab but at a different time. The videos were displayed using E-Prime (EPrime, 2015) for randomisation and data collection on a standard desktop computer with inbuilt speakers for audio. A continuous wave functional near infrared spectroscopy (fNIRS) system (fNIRDevices, 2016) was utilised to measure prefrontal cortex activity. fNIRS is similar to functional magnetic resonance imaging in that it provides a hemodynamic measurement of brain activity. fNIRS employs optical signalling to record oxygenated and deoxygenated haemoglobin concentration changes which, in turn, can be

used to provide a measure of cognitive workload. By recording optical density changes at two differing wavelengths (700nm and 830nm) the relative change over time in oxygenated and deoxygenated haemoglobin can be calculated using the modified Beer-Lambert Law. Data was collected at 2Hz using Cognitive Optical Brain Imaging (COBI) Studio software (COBI, 2015). Raw fNIRS light intensities were subjected to signal processing to remove any motion or physiological artefacts (heart rate or respiration rate) by employing a linear phase low pass finite impulse response filter with a cut off frequency of 0.1Hz.



Fig. 24. Participant environment. The participant is sat facing a desktop machine that is running the sender movies in E Prime. The participant is wearing an fNIRS sensor with a scarf wrapped around it to ensure light did not affect the data captured.

4.4.4 Variables

The dependent and independent variables are described below.

Independent Variables

The conditions looked at were medium and truthfulness. The medium of video conferencing was tested against that of virtuality telepresence. Two senders were each asked to give both a true and false account in answer to four questions.

	Videoconferencing	Virtuality Telepresence
Truthful account	VCt	VTt
Deceitful account	VCd	VTd

Table 5. Independent variables

Dependent Variables

The dependent variables were the success of determining truthfulness in each medium and changes in the level of cognitive effort while doing so. The scoring scheme was simple in that the correct identification of truthfulness (i.e. the correct identification of a truthful or a deceitful account) was scored with 1 and an incorrect judgement with 0. Thus, for a given medium, given that there were 6 accounts, if all truths and lies were correctly identified a maximum score of 6 would be obtained, and if none were identified the participant would obtain a total score of 0. In addition, data on cognitive load using fNIRS were collected while the receivers conducted the task. This data measured the change in the oxygenation of the blood in the prefrontal cortex. Increases in oxygenation are an indication of an increase in cognitive load.

	Description	Range	Type
Score	Number of successes in determining the truthfulness of account	0 to 6	Task performance
Oxygenation	Change in oxygenation of the blood in the prefrontal cortex	-1 to 5	Difficulty

Table 6. Dependent Variables

4.4.5 Ethics

Ethical approval was gained for the study from the University ethics board. The approval number was CST 12/51. The ethical approval included approval of the information sheet and the consent form. The information sheet explained to the participants the general theme of the study and what they might expect. The consent form informed participants of the anonymity of data, the right to leave at any time and the right to have one's data completely removed.

4.4.6 Participants

In total N=23 participants were involved in the experiment. Two confederates were initially captured to produce the test media. Twenty one participants undertook the lie detection experiment.

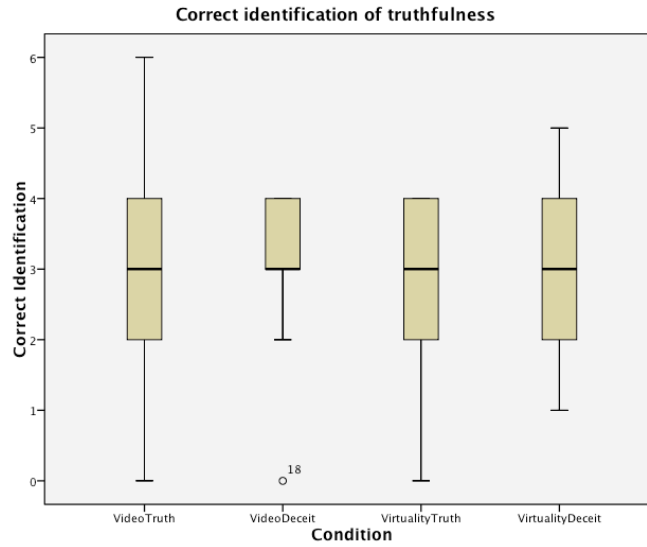
4.4.7 Analysis method

Paired samples t-tests were used to compare the correct identification of truthfulness. A Pearson correlation test was used to compare cognitive load.

4.5 Results

A quantitative analysis looked at task performance and an indication of cognitive load, and then correlated the two.

When testing H1 (Accuracy of determining truthfulness will be significantly less when determining truthfulness through the reconstructed avatar), no significant difference was found. This hypothesis was tested using a paired samples t-test and rejected ($t= 1.072$, $p=0.297$, $df=20$), see Graph 5 below.

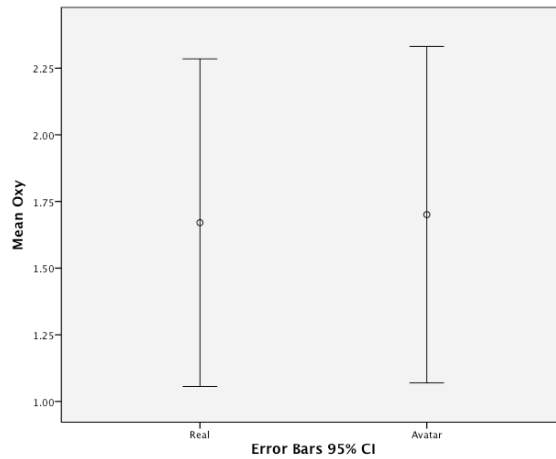


Graph 5. Task performance - correct identifications of “truthfulness”

Real		Reconstruction	
Truth	Deceit	Truth	Deceit
66	64	58	63

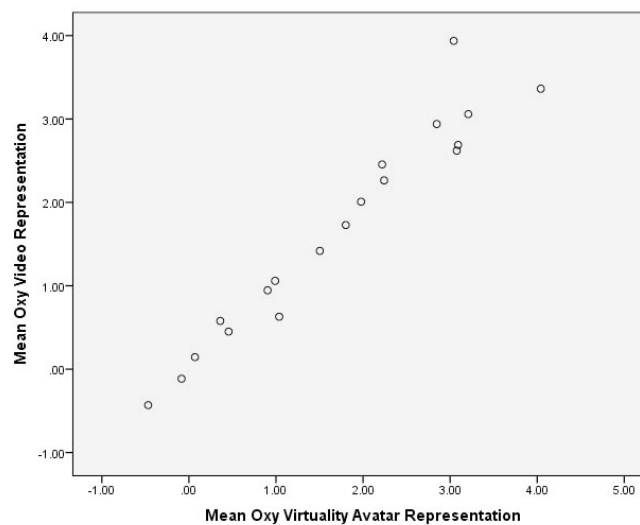
Table 7. A table showing the correct identification of truth and deceit. The totals are out of a possible 126 statements.

When testing H2 (Cognitive workload will be significantly higher when determining truthfulness through the reconstructed avatar), the result was that cognitive load, as indicated by change in the oxygenation of blood in the frontal cortex, under the two conditions, was not found to be significant ($t = -0.40$, $df = 18$, $p = 0.694$). See Graph 6 below.



Graph 6. Comparing change in cognitive load across the two mediums - mean change in the oxygenation of blood in the prefrontal cortex for two groups

An interesting follow up to this finding is that when the two variables are correlated using the Pearson product-moment correlation coefficient the finding is highly significant ($r=0.968$, $N=19$, $p<0.01$). This finding strongly suggests that receivers who found the task difficult when looking at the video had a similar difficulty when looking at the virtuality telepresence version, and vice versa. This would indicate that the virtuality telepresence version is a very good virtual equivalent of the video version.



Graph 7. Mean change in oxygenation when viewing video representation against mean change in oxygenation when viewing virtuality avatar representation during the detection of truths.

4.6 Discussion

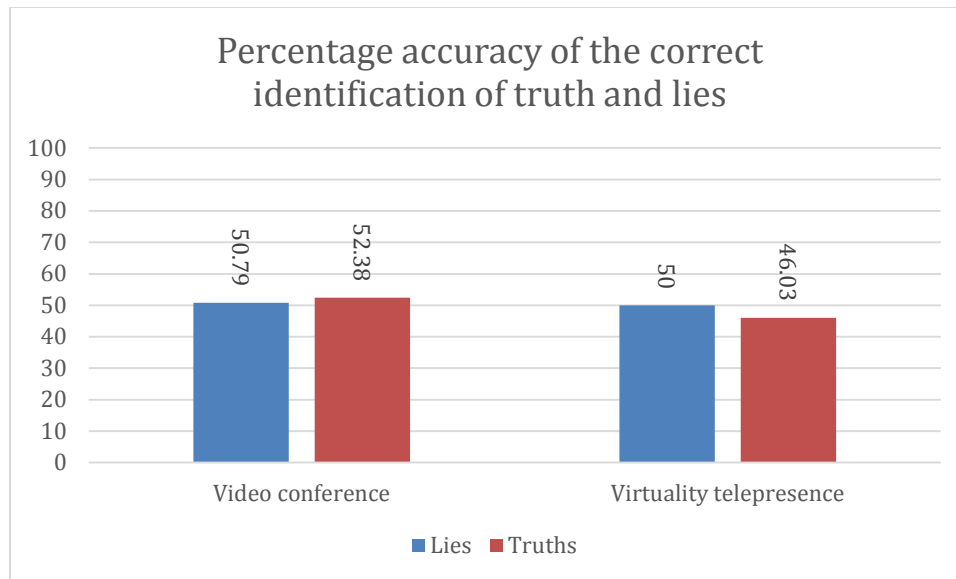
This experiment set out to determine whether 3D reconstructed avatars created using shape-from-silhouette displayed the same truth bias that authored abstract 3D avatars were shown to in a comparison of a small number of previous studies. The researcher suggests that talking to a user wearing an abstract 3D avatar, with no facial animation, is similar to talking to somebody wearing a mask. Therefore the research questions were firstly: Would a 3D medium that removed this mask, result in a truth bias akin to video and thus the real world? Secondly, do the shortcomings of the tested method for reconstruction contribute to a distortion in truth bias? In order to test these questions the virtuality telepresence medium first needed to be tested independently. This was achieved by comparing the virtuality telepresence condition against that of video conferencing, testing whether the visual quality of the medium would impact upon a receiver's ability to detect truthfulness. The results from this experiment would then allow a comparison of the medium with those tested in previous studies from both VR and psychology. The following hypotheses were tested and the outcome for each hypothesis is presented.

H1: The accuracy of determining truthfulness will be significantly less when viewing the reconstructed avatar – Overall there was no significant difference in accuracy between video conferencing and virtuality telepresence.

H2: Cognitive workload will be significantly higher when determining truthfulness when viewing the reconstructed avatar - There was found to be no significant difference in levels of cognitive effort when identifying truthfulness across media. There was, however, a

significant correlation between workload across each medium, suggesting that receivers have to work equally hard when viewing either medium.

No significant difference was found in the accuracy of determining truthfulness across the conditions ($t= 1.072$, $p=0.297$, $df=20$). This was the case when the sender was telling the truth or lies, and overall. There was no significant difference in the number of accounts considered to be true or false ($t= 1.072$, $p=0.297$, $df=20$). This suggests that there was no difference in truth bias across the conditions. These results are consistent with other studies using videos but different to those using CGI avatars (with limited facial animation) or voice alone. However, a negative truth bias was evident between the 3D video based medium and previous psychological studies. The significance of the latter was not testable, due to the detail of reporting in the literature. There was a significant correlation in changes in cognitive workload between conditions ($r=0.968$, $N=19$, $p<0.01$). This suggests that individuals worked equally hard to determine the truthfulness from the footage of both mediums. This suggests that the 3D medium did display a sufficient quality of representation for observers to determine the truthfulness at a similar level as that for video conferencing.



Graph 8. Accuracy of determining truthfulness across the media

Graph 8 visualises the accuracy of determining truthfulness, for the true statements and the false statements. Both mediums have an average overall accuracy close to chance (50%). Video conferencing showed a combined average of 51.58% and virtuality telepresence 48%. These are close to the 54% average of the real world studies reported by Bond and DePaulo (CF Jr Bond, 2006). The reported differences in accuracy when identifying truth and lies is also small. It is much smaller than those shown in the avatar study reported by Steptoe et al. (Steptoe et al., 2010) which lacked facial expressions. The avatar study by McDonnell et al. (McDonnell et al., 2012) that did reproduce facial expressions, reported a truth bias close to real world studies. It should be noted that, due to the small number of studies available for comparison and the relatively small number of participants when compared to real world studies, a much larger study testing all conditions would be needed to draw any firm conclusions. The above results were promising for the medium of virtuality telepresence. They show that, even with the impacting factors of shape-from-silhouette,

virtuality avatars contain enough detail to allow the receiver to ascertain truthfulness comparably with video conferencing.

4.7 Summary

This experiment set out to compare the detection of deceit across 2D and 3D mediums which were designed to faithfully represent appearance and identity. The viewpoint of the latter was fixed to isolate the impact of the quality of facial representation from mutual gaze. A free viewpoint would have allowed the user to view the avatar from any angle making a direct comparison of quality impossible. In addition, it would have added an interaction task that would have impacted on the cognitive workload of the participant. The first contribution is to demonstrate, through an experiment, that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of deceit. The second contribution is to demonstrate through an experiment that participants work equally hard when detecting deceit through a 3D reconstructed medium and video conferencing. The third contribution is to link the two to provide additional evidence to support the above trend.

In addition, there are three methodological contributions:

1. Locking and matching a viewpoint to allow exact contrast of visual quality without the distraction of interaction.
2. An experimental setup for simultaneously capturing 2D video and video based 3D reconstructed video for contrast purposes.
3. The use of fNIRS to compare cognitive effort during the detection of truths in 2D video and video based 3D reconstruction for contrast purposes.

This experiment has shown that a 3D graphics avatar, continually reconstructed from live video, is comparable to video conferencing in portraying the overall truthfulness of the person it mimics. A comparison with the small amount of previous literature on this subject suggests that people trust others too much when a medium does not reproduce their facial expression; however, a much larger study would be needed to draw any firm conclusions. In addition, it would be interesting to examine the impact of a free viewpoint and to explore how the ability to move around the sender would impact upon how well the receiver is able to detect deception and identify body language, thus testing whether freedom of movement and ability to interact is a help or a hindrance.

These findings may contribute to the design of future virtual humans and communication mediums. They can also be relevant regarding the choice of mediums used in the justice system, in advertising, social media, politics, distributed family and the workplace. Not everyone will want to remove the mask but perhaps further research can inform others of its influence.

5 Chapter 5 - Experiment 3: Facial Action Coding – a comparison of video frames and 3D video frames in terms of the quality of facial features and expressions.

5.1 Introduction

This experiment measured the quality of facial representation afforded by video based 3D reconstruction. This was achieved by comparing still frames from video conferencing and a 3D video based computer graphic medium. The purpose was to determine if the process of 3D reconstruction reduced the detail of facial expressions. N=14 separate participants (coders), with a psychology background, were asked to identify facial muscle movements. Data were captured using facial action coding. Results show that, despite the obvious visual reduction in quality and the resultant loss of some facial detail, receivers performed and worked to a similar degree in both mediums. The findings are of interest to those developing 3D telepresence technologies and virtual humans.

5.2 Method

5.2.1 Identification of facial expressions

This section outlines the design and implementation of a facial expression study. The aim was to understand whether facial muscle movements were discernible in each medium. Specifically, the study shows synchronised pairs of images from the two mediums. The frames used were picked as they show interesting facial expressions. The frames shown below (see Fig. 25) were representative of the clarity of representation of these facial features across all the videos.



Fig. 25. Still frames showing video and virtuality avatar stimulus images.

5.2.2 Variables

The independent and dependent variables are now discussed.

Independent variable

The independent variable is the visual medium consisting of video conferencing and virtuality telepresence.

Dependant variable

The dependant variables in this study were the facial action coding units (Ekman and Friesen, 1978) available to the coders. The scoring scheme was simple in that the correct identification of an action unit was scored with a 1. The following facial action units were tested and are summarized as given below.

Upper face: 1, 2, 4, 5, 6, 7, 43, 45, 70, 71. Head Position: 51, 52, 53, 54, 55, 56, 57, 58.

Eye Position: 61, 62, 63, 64, 65, 66. Lips Parting Jaws Opening: 25, 26, 27. Lower Face: 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 23, 24, 28, 72. Miscellaneous action units: 8, 19, 21, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39

5.2.3 Procedure

Fourteen participants were asked to score a set of ten images. They were split into two groups, one group scoring the video conferencing version of the images, the other the virtuality telepresence images. They were given scoring sheets containing all the facial action units and a presentation on how the system works. They were then asked to score a series of images. The experiment lasted approximately 20 minutes and the participants were not paid for taking part.

5.2.4 Subjects

A group of fourteen Psychology undergraduates participated in the study. They had not seen the images before. They were given a short presentation on how to code the images prior to the task.

5.2.5 Participant environment

A computer lab with one central bench and two sets of PCs running down it on either side was used to display the still images. Half of the students faced each set of PCs. One set of students marked the video conferencing stills and the other the virtuality telepresence stills. They could not see the other groups' images.

5.2.6 Analysis

Related t-tests were used to compare the identification of facial action units.

5.3 Results

A qualitative analysis compared the synchronised images from both mediums in terms of facial qualities using the Facial Action Coding System (FACS).

When testing H1 (Significantly fewer facial actions will be detected through the reconstructed avatar), no significant difference was found overall. However, over both mediums, coders identified a total of 626 facial action units with a mean of 4.47 facial action units per image. The video conferencing condition had 339 facial action units (mean 4.84

per image) and the virtuality telepresence condition 287 (mean 4.1 per image). There was no significant difference between the overall individual action units identified across each condition (video conferencing mean = 5.84, virtuality telepresence mean = 4.95, $t=1.519$, $df=58$, $p=0.134$). However, there was a significant difference between the lower facial action units identified (video conferencing mean = 7.0, virtuality telepresence mean = 4.19, $t=2.858$, $df=15$, $p=0.012$). The upper face action units and head position had identical scores.

5.4 Discussion

This experiment set out to determine the quality of 3D reconstructed avatars created using shape-from-silhouette. The question was asked: Do the shortcomings in the method for reconstruction contribute to a reduction in facial detail? An answer to this question was achieved by comparing the virtuality telepresence condition against that of video conferencing thus testing whether the visual quality of the medium would impact upon a participant's ability to detect facial muscle movements. The following hypothesis was tested:

H1: Significantly fewer facial actions will be detected when viewing the reconstructed avatar. The experiment examined which parts of the face the receiver could identify using facial action coding. Ten still frames from each of the mediums were scored by a group of 14 coders. The results showed no significant difference in the overall number of facial action coding units identified across video conferencing and virtuality telepresence. This suggests that, even with the known problems of shape-from-silhouette, the image quality was good enough for identifying facial movements on a par with video conferencing.



Fig. 26. Screen shot of virtuality avatar illustrating: 1. Slicing. 2. Lack of Concavities. 3. The droop effect.

The above results were promising for the medium of virtuality telepresence. They show that, even with the impacting factors of shape-from-silhouette, virtuality avatars contain enough detail to allow a participant to identify facial muscle movements comparably with video conferencing. One problem with the fundamentals of the approach was shape-from-silhouette's inability to detect concavities, although the general 3D geometry and texturing that illustrated the concavities contained enough information to allow the receiver to identify them (see Fig. 26). The other issues were around camera setup with spatial calibration causing slicing (see Fig. 26). In this experiment the slicing did not appear to affect performance. However, it is noted that if the slicing had occurred in a different location, at the jaw perhaps, then this may have been more of an issue. It is possible that droop caused by camera position could have been an issue with a significant difference in the number of lower face action units identified across the medium. This could also have been affected by one of the participants having a beard that might have made identification more difficult, however, in such a case this would have been present over both mediums. Finally, there was a problem with colour calibration causing an unrealistic skin tone although this did not appear to affect the results.

5.5 Summary

This experiment measured the quality of facial representation afforded by video based 3D reconstruction. This was achieved by comparing between still frames from video conferencing and a 3D video based computer graphic medium. The contribution of this work is, firstly, to demonstrate that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of facial muscle movements and, secondly, it is a methodological contribution showing how facial action coding can be used as a measure of quality for virtual character facial expressions. Overall, there was no significant difference in the number of facial actions identified by coders across the two mediums (video conferencing mean = 5.84, virtuality telepresence mean = 4.95, $t=1.519$, $df=58$, $p=0.134$). This suggests that the two mediums are equally effective at displaying facial muscle movements. However, there was a significant difference across conditions for the lower facial action units (video conferencing mean = 7.0, virtuality telepresence mean = 4.19, $t=2.858$, $df=15$, $p=0.012$).

This study has shown that a 3D graphics avatar, continually reconstructed from live video, is comparable to video conferencing in portraying the overall facial muscle movements of the person it mimics. These findings may contribute to the design of future virtual humans and communication mediums.

6 Chapter 6 – Discussion

6.1 Overview

The portrayal of behaviour by virtual characters is an enormous area, far too large to cover in one thesis. The challenge outlined at the start of this work was to explore how visual qualities impact on behavioural realism and interpretation of meaning, with the overarching research problem being: how is the identification and interpretation of behaviour impacted upon by differences in visual and behavioural realism across 2D video and 3D media? The findings of the research are discussed below to illustrate how the above has been addressed.

Traditionally, when creating virtual characters, developers needed to ensure visual and behavioural quality high enough for the identification of basic movements, such as walking, running and jumping. However, in virtual reality and serious gaming applications, the interpretation of behavioural meaning can be more important than such interpretation within entertainment titles. In this thesis the focus was not only on the identification of behaviour, but also on its underlying meaning. It might be possible to identify a walk or other movement, but what about the intent and meaning of the behaviour? In this thesis the effect of visual and behavioural realism on the identification and interpretation of NVB in virtual characters has been examined, and each experiment compared a 3D medium to 2D video, locking and matching a viewpoint to allow an exact contrast of visual quality without the distraction of interaction.

The first experiment was a study of gross NVB identification in primates. The second experiment studied more subtle behaviours and our interpretation of them. This was achieved by examining the identification of truths and the cognitive effort required in doing

so. The final experiment explored the identification of facial muscle movements in humans. The combination of surveying the literature and conducting these studies allowed an examination of virtual characters' ability to display NVB with enough detail to allow identification and intent to be distinguished.

The aim of this PhD was to investigate how the realism of virtual characters affects the identification of behaviour and our interpretation of it. This has implications for applications, ranging from social media to serious training (Bailenson et al., 2006b). Avatars, worn by a user, will have an impact on the user's behaviour within the virtual world (Peña et al., 2012) (Fox et al., 2013) and influence other users (Behrend et al., 2012) possibly also temporarily affecting behaviour outside the VE (Rosenberg et al., 2013). Avatar appearance can influence our perception of the character's empathy towards us (Dean et al., 2009) and our self-disclosure to them (Bailenson et al., 2006b). In the real world, a person's appearance can impact upon another's trust of them (Spellman and Tenney, 2010) suggesting that it could be useful to study how an authored abstract avatar's appearance can impact upon trust. Avatars enable users to do something that they cannot do in reality, in that they can easily change their face, size and shape (Bailenson et al., 2004, Bailenson and Beall, 2006). In reality this cannot be done unless through having an operation or putting on a mask. This is of particular relevance to serious applications. The correct type of avatar, for example, can give the wearer an advantage in a job interview (Behrend et al., 2012). This is of great benefit for the wearer of the avatar but not necessarily for the viewers of it, who may have been duped into believing someone is professional or looks in a particular way. The ability of an avatar to change a person's perception is similar to how our real life perception of each other can change according to appearance.

Whatever form a character takes it is important to ensure that the quality of visual appearance and behavioural representation are high enough to make it fit for purpose. If the character is going to be a Chimpanzee in order to teach primate behaviour to students, then the behavioural representation needs to be high enough to illustrate not just a movement but also the intention behind it. The same applies to virtual humans; subtle behaviours can give clues concerning intentions and, if characters are to be used as avatars in more serious social interactions, then the relevant subtleties should be represented at an adequate level.

6.2 Contributions

6.2.1 Contributions from the Literature Review

1. The first contribution was the undertaking of a new literature survey identifying the following gaps in knowledge:

- i. The identification of a possible trend/opportunity across the literature that suggests that zoos need a better way of informing visitors about animal behaviour; this ‘better way’ can contribute to serious gaming and conservation education.

This contribution concerning the area of serious games identifies a possible trend/opportunity across the literature that suggests that zoos need a better way of informing visitors about animal behaviour which would also contribute to conservation education. This led to the first experiment of the thesis which examined the identification and interpretation of Chimpanzee behaviour. The experiment demonstrated that, overall, there was no significant difference between video and CGI for this task.

- ii. The identification of a possible gap and opportunity in the literature that suggests zoologists would benefit from a better way of learning about animal behaviour.

The literature identifies that zoologists sometimes mimic animal behaviour when caring for them (Jenvoid, 2008, Case et al., 2015). Jenvoid suggests that Chimpanzees responded in a more playful way when keepers interacted with them using Chimpanzee behaviour (Jenvoid, 2008) while, in a past study, Babu et al. demonstrate that VR can be effectively used to teach human non-verbal behaviour (Babu et al., 2007). Currently behaviour is learnt from the observation of the real world, from video or books. Serious gaming and VR offer the ability to replay behaviours over and over again and, potentially, as shown by Babu et al. (Babu et al., 2007), offer a way to validate their accuracy. Serious gaming and VR can also allow training in the tasks that may be too expensive or dangerous to be conducted in real life (Ieronutti and Chittaro, 2007, Falloon, 2010, Dalgarno and Lee, 2010). The use of VR would allow zoologists to benefit from a better way of learning about animal behaviour and, in particular, how to mimic it.

- iii. The identification of a possible trend across the literature that suggests reproducing facial expressions reduces the tendency of observers to over trust. However, this study has identified only a limited number of studies in this area, so more work would be needed to confirm this.

This finding contributes towards the area of virtual characters by identifying a possible trend across the literature that suggests reproducing facial expressions reduces the tendency of observers to over trust. It should be noted, however, that due to the small number of studies

available, further work would be needed to confirm this. There is already literature from real world (CF Jr Bond, 2006) and authored abstract avatar study's (Steptoe et al., 2010, McDonnell et al., 2012); however, nothing yet existed for the field of video based 3D reconstructed avatars. It was, therefore, interesting to test this medium alongside others and see if a truth bias was present. The presence of truth bias would be important because it would allow users to pick an avatar with a lower behavioural realism in order to gain a trust advantage. This could be in addition to other advantages that could be gained through the character's visual appearance. In addition, the study has identified that facial action coding can be used as a way of comparing quality across different mediums.

6.2.2 Contributions from Experiment 1: Gross behaviour identification – a comparison of video and CGI for the identification of Chimpanzee behaviour.

Primary Contributions

- 1. The first primary contribution is demonstrating that CGI with low level NVB is sufficient for the identification of gross NVB on a par with video.**

The first experiment of the thesis examined the ability of CGI Chimpanzees to convey gross non-verbal behaviour. The results showed that there was no significant difference between CGI and video in this task. Overall, real video (83.1% correct) performed better than CGI (79.4% correct) but not significantly ($p=0.377$). This result paves the way for further studies on interaction and highlights the use of serious games as a possible way of teaching zoo staff and visitors about animal behaviour in a more interactive way. This would be beneficial as previous literature suggests visitors learn more from interpretive presentations (Visscher et

al., 2009). However, such interpretive presentations in zoos only happen at certain times; serious gaming offers a way to allow instant access to such presentations, contributing to conservation education which is one of the main aims of Zoos (WAZA, 2005).

2. The second primary contribution is identifying a possible issue with the identification of behavioural intent due to a lack of subtle NVB.

Participants suggested that the CGI Chimpanzees lacked the ability to convey subtle movements and intent. For example, yes, a chimpanzee was approaching another but why? “Key issue is granularity. Gross actions were identifiable in both, but nuance and purpose only really came through in the film. For approach you could see contact between two individuals in film but VR (CGI Film) did not establish this and the animals could have just wandered closer.” This is important, and links well with the identification of a need for a behavioural trainer for zoologists. It illustrates that, although identification of behaviour is possible with a lower level of realism, more subtle behaviours are necessary to allow the understanding of intent. This fact contributes to knowledge to this area.

Methodological Contribution

1. Motion tracking a zoo keeper to mimic primate behaviour.

Placing tracking markers on wild animals is not ethical, nor would it result in natural behaviour. Putting animals into a large capture volume is also not plausible as, again, behaviour would not be natural due to the new surroundings. Inspiration was taken from films where actors are used for motion capture. However, in this case a Chimpanzee keeper

was used to mimic the animal's movements and behaviours, thus contributing a novel solution to behaviour representation.

6.2.3 Contributions from Experiment 2: Lie detection – a comparison of video and 3D video for the detection of deception.

This contribution is made through comparing the qualities of 2D video and video based 3D reconstructed avatars when identifying truths, exploring how visual qualities impact upon trust during lie detection and exploring also the cognitive effort while doing so. Measures included the participant's cognitive workload and accuracy of judgments. Comparison was also made to a small number of previous VR studies suggesting that appearance may impact upon how avatars are perceived.

Primary contributions

- 1. The first primary contribution is demonstrating through an experiment that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of deceit.**

The experiment conducted compared lie detection across the mediums of 2D video and video based 3D reconstruction, then discussed the results alongside existing lie detection studies in both the real world and VR. A small number of previous studies suggest a truth bias exists when facial details/expression are limited. The experiment examined whether this bias existed with the medium of video based 3D reconstruction, or whether it was more akin to standard video and, therefore, real life.

No significant difference was found in the accuracy of determining truthfulness across conditions ($t= 1.072$, $p=0.297$, $df=20$). This was the case when the sender was telling the truth or lies, and overall. There was no significant difference in the number of accounts considered to be true or false ($t= 1.072$, $p=0.297$, $df=20$). This suggests that there was no difference in truth bias across the conditions. These results are similar to other studies using videos but are different to those using CGI avatars (with limited facial animation) or voice alone. However, a negative truth bias was evident between the experiment's 3D video based medium and previous psychological data. The significance of the latter was not testable, due to the lack of details in the reporting in the literature.

This contributes knowledge to the area of virtual characters; in particular, regarding their ability to faithfully portray truths. There are now studies for real world, for authored abstract avatars with no facial muscle animation, with facial muscle animation and for virtuality avatars (provided in this thesis). However, due to the small number of studies available it would be necessary to conduct a larger study covering all simultaneously, in order to draw a firm conclusion. However, as previously discussed, hardware issues currently stop this from being possible when using the same data set.

2. The second primary contribution is demonstrating through an experiment that participants work equally hard when detecting deceit through a 3D reconstructed medium and through video conferencing.

There was a significant correlation in changes in cognitive workload between conditions ($r=0.968$, $N=19$, $p<0.01$). This suggests that individuals work equally hard to determine truthfulness from the footage of both mediums. Thus this is a contribution to the area of

virtual characters demonstrating fNIRS use a measure of user effort during the detection of truths.

3. The third primary contribution is linking the two previous contributions to provide additional evidence to support the above trend.

The results of this study illustrate that there is no significant difference between video and video based 3D reconstruction for the detecting of truths. Moreover, changes in participants' cognitive workload during the task suggest that they work equally hard when looking at each medium, pointing to equal performance and effort when compared to video.

Methodological Contributions

1. Locking and matching a viewpoint to allow exact contrast of visual quality without the distraction of interaction.

The viewpoint of the video based 3D reconstructed avatar was fixed to isolate the impact of the quality of facial representation from mutual gaze. A free viewpoint would have allowed the user to view the avatar from any angle making a direct comparison of quality impossible. In addition, it would have added an interaction task that would have impacted on the cognitive workload of the participant.

- 2. An experimental setup for simultaneously capturing 2D video and video based 3D reconstructed video for contrast purposes.**

A novel experimental setup allowed the simultaneous capture of the video based 3D reconstructed avatar and the teleconferencing medium.

- 3. The use of fNIRS to compare cognitive effort during the detection of truths in 2D video and video based 3D reconstruction for contrast purposes.**

The use of fNIRS allowed an examination as to whether rendering artefacts from the method of reconstruction forced the participant to work harder during identification, thus allowing a physiological measure of quality.

6.2.4 Contributions from Experiment 3: Facial Action Coding – a comparison of video frames and 3D video frames for quality of facial features and expressions.

Primary Contribution

- 1. The primary contribution of this work is demonstrating that a 3D medium aimed at faithfully reproducing a person's appearance can be comparable to video conferencing in the overall portrayal of facial muscle movements**

This contribution is made by examining how the visual qualities of video based 3D reconstruction and 2D video affect the recognition of human facial expressions. This was undertaken by comparing each medium's ability to convey facial muscle movements, and this was measured using facial action coding. The first study examined the use of abstract 3D avatars of chimpanzees for gross behaviour identification. One point raised by the

participants was the lack of subtle behaviours that allow intent to be identified, meaning that behaviours could be identified but the meaning behind them is somewhat lost. It was, therefore, decided to focus on one such subtle behaviour, in this case facial expressions.

Ideally, the study would have covered authored abstract and video based 3D reconstructed avatars as both allow the display of facial expressions. Software such as faceshift (Faceshift, 2015) allows facial expressions to be faithfully applied to abstract avatars within game engines in real time, while video based 3D reconstruction allows participants to be 3D scanned and faithfully reconstructed in real time, capturing facial expressions and subtle body movements in the same process. However, due to hardware constraints previously mentioned, it was not possible to simultaneously capture video, abstract avatar and video based 3D reconstruction simultaneously, thus 2D video and video based 3D reconstruction were contrasted to check whether the latter conveyed enough visual detail for the correct identification of facial muscle movements. This was necessary as 3D reconstruction can sometimes contain rendering artefacts such as slicing or droop. Slicing, depending on location, can cause detail to disappear. Droop can make a character look older, unwell or sad (Roberts et al., 2015a) and is caused by camera position. Additionally, texturing and image contrast errors can cause striping. The results showed that there was no significant difference in facial action units identified overall. However, on closer inspection, there was a significant difference between the lower facial action units identified. This could have been a direct result of droop as the camera positions in the capture system are based above the user. In summary, overall, there was no significant difference in the number of facial actions identified by the participants across the two mediums ($t=1.519$, $df=58$, $p=0.134$). This suggests that the two mediums are equally effective at displaying facial features.

However, there was a significant difference across conditions for the lower facial action units ($t=2.858$, $df=15$, $p=0.012$).

The results of this experiment illustrate that video based 3D reconstruction is, indeed, capable of the successful illustration of facial muscle movements and expressions. However, there was an issue with recognition of the lower facial action units, possibly caused by droop which, in turn, points to the camera placement not being as good as it could be. The results illustrate that video based 3D reconstruction, even in its present form, can adequately display facial muscle movements. This would suggest that the medium, given better camera alignment and a high enough frame rate, would support both facial and micro expressions in real-time. The experiment also demonstrates how facial action coding can be used as a measure of quality when comparing virtual characters' ability to portray subtle facial expressions, thus contributing to the field of virtual characters.

Methodological Contribution

1. The use of facial action coding to measure the quality of video based 3D reconstruction when compared to 2D video.

The use of FACS gives a testable set of variables to directly contrast quality across media. If one does not show the same level of detail then the visual display of facial movements will be lower, therefore, affording a lower level of behavioural representation for facial expressions.

6.3 *Critical evaluation*

6.3.1 Literature review

Due to the research area being multidisciplinary, it required the literature review to cover a large variety of subjects. This review was fascinating but it did not allow the same level of rigour as a review focusing solely on one topic. The review was also impacted upon by the change in direction from animals to humans which, although interesting, added further work. However, changes in research focus are common as new and exciting areas open up during the course of study. In addition, looking at a wide spread of literature is a necessity for this kind of study in order to encapsulate all the areas needed for the research.

6.3.2 Experiment 1: Gross behaviour identification – a comparison of video and CGI for the identification of Chimpanzee behaviour.

There are two reasons why a comparison of CGI and video for identification of behaviour is relevant. Firstly, observations of human or animal behaviour are mainly conducted in person or by watching video recordings. Virtual environments offer a free viewpoint update as in real life. A fair comparison of the qualities of virtual environments and real world for the identification of behaviour is not possible as people and animals do not display behaviour on command, therefore, the media cannot be matched. In order to overcome this issue, video observations were recorded and equivalent CGI footage created. This separation of the interactive experience allows the visual medium to be tested in isolation removing any confusion in future studies as to whether the results obtained are due to interactive or visual differences, thus allowing a fair isolated comparison of visual media. This is a similar

approach to that of Ragan et al. (Ragan et al., 2012) who are also studying the impact of individual qualities of virtual environments.

Users did perform well when identifying behaviours so there was enough realism for the task set out. However, for a serious training application, such a level of realism as that given in this study may not be enough; it should be possible to identify an aggressive walk or subtle facial expressions that can precede displays of violence. Simply put, users need to be able to understand intent. For obvious reasons it was not possible to facially track Chimpanzees. It is, however, possible to track a human and if it was proved feasible to reproduce subtle human behaviours then other primate behaviours would also be viable.

The experiment required participants to fill out a questionnaire declaring whether they had identified a particular behaviour. With hindsight, it would have been useful to record exact numbers either by using a clicker or by an audio recording of participants. It would also have been useful to conduct the study a second time with the same participants at a later date. This would have allowed for the study of knowledge retention and the learning of behaviours in addition to identification.

6.3.3 Experiment 2: Lie detection – a comparison of video and 3D video for the detection of deception.

In an ideal scenario it would have been preferable to capture video, authored abstract and virtuality representations of the participant (sender) simultaneously. However, hardware issues (such as occlusion) prevented this from happening. This would have allowed a direct comparison of the same data set. The comparisons provided are with a small number of past studies which would have had varying experimental designs and setups. However, the study

is still able to highlight what appears to be a suggestion that avatars lacking facial expression display a considerably higher truth bias than normal. Further studies are needed in this area to draw firm conclusions.

6.3.4 Experiment 3: Facial Action Coding – a comparison of video frames and 3D video frames regarding the quality of facial features and expressions.

The original aim was to compare video, abstract avatars and video based 3D reconstructed avatars. However, as previously discussed, facial capture for abstract avatars requires a camera to be close enough to the face to capture facial movements. This would have occluded facial capture using video based 3D reconstruction and, therefore, it was decided to compare video and the relatively new medium of 3D reconstruction as both could be captured simultaneously using a new experimental design that allowed timings and quality to be matched to allow a fair comparison. This does mean, however, that a further study will be necessary to identify whether current methods of mapping facial muscle movements to authored abstract avatars are sufficient to adequately represent behaviour. It would also be desirable to put a much larger number of images through this process and perhaps go to a more granular level in order to identify exactly which points of the face appear to be most affected by reconstruction. However, the study does highlight the lower face as being particularly problematic which fits in with the known problem of droop and, overall, the experiment shows no significant difference between the numbers of facial action units identified.

7 Chapter 7 - Conclusions and future work

The overarching research question of this thesis was:

How is the identification and interpretation of behaviour impacted upon by differences in visual and behavioural realism across 2D video and 3D media?

Knowledge of social interactions has traditionally been passed on through observation and standard teaching materials such as lectures, books and documentaries. VR and serious gaming have presented opportunities to explore new and novel ways of learning about behaviour where the user can be transported into a virtual world and learn through simulation and ‘doing’. In this thesis the focus was on the identification and interpretation of non-verbal behaviour, ranging from gross full body behaviour to subtle facial expressions. The use of virtual characters as avatars in games, social media and serious applications has become increasingly popular. It is, therefore, important that there is an understanding as to how our and others’ choice of avatar impacts upon the identification and interpretation of non-verbal behaviour. The ability to identify virtual character behaviour, and interpret it, is paramount in the usage of avatars representing humans in serious applications, telepresence or entertainment.

This study successfully illustrated how abstract virtual characters can be used to convey gross full body behaviour, and identified that there is a lack of detection of intent. For this reason, attention was then shifted to more subtle facial expressions. Due to the problems previously discussed, the capture of subtle behaviours from animals is extremely challenging, therefore, attention switched to human behaviour. This was undertaken by comparing video based 3D reconstruction with video while detecting truths (in which there

was no significant difference between the mediums and, in addition, participants appeared to work equally hard when conducting this task in both mediums).

When the results were compared to the small number of previous studies available, it was discovered that avatars lacking facial expressions appear to display a truth bias. However, it should be noted that experimental design of each of these studies, including this one, was not identical. Thus, as previously stated, a much larger study would be required to draw any firm conclusions. Subsequently, the next stage in the research was proving that video based 3D reconstructed avatars can display facial muscle movements on a par with video, with the exception of lower facial action units (which, it is suggested, was due to the droop effect caused by camera positioning).

Overall, the thesis shows how both authored abstract and 3D reconstructed avatars can be used to display both gross and subtle NVB on a par with video, through both experimentation and a comparison with the literature. It is suggested that, currently, most avatars lack enough subtle behaviour to allow the identification of intent. Time, cost and hardware are all factors which affect research into this area. Both authored abstract and video based 3D reconstructed avatars have the ability to demonstrate NVB. It is suggested that serious applications require higher levels of behavioural realism to reach their full potential, and that most VR or gaming applications would benefit from planning greater levels of character relevant realism when planning agents or avatars, therefore, making them more believable and fit for purpose.

The studies identified that, when presented through a fixed viewpoint, 3D characters (authored abstract or reconstruction) can perform on a par with video to display behaviour, thus allowing for behaviour identification and interpretation, as indicated below.

Abstract avatars versus video – identification of gross non-verbal behaviour.

3D reconstruction versus video – identification of subtle non-verbal behaviour (facial muscle movements) and the ability to portray truth.

Future studies could now build upon this thesis by freeing the user viewpoint and conducting similar studies to identify whether the addition of interaction and movement is a help or a hindrance. In addition, it is suggested that further studies should be conducted examining virtual character visual and behavioural realism within virtual environments and the relationship with user trust and perception. If such research can be used to gain any advantage then it would have implications for many applications such as courtroom interviews, social media and gaming. If a real person showed little or no non-verbal behaviour when interacting then people would think twice about trusting him/her. This should be the same in the virtual world. This could become an exciting area of research in the near future, as avatar mediated communication becomes more widely adopted.

List of references

- ABIRACHED, B., ZHANG, Y., AGGARWAL, J. K., TAMERSOY, B., FERNANDES, T., MIRANDA, J. & ORVALHO, V. Improving communication skills of children with ASDs through interaction with virtual characters. *Serious Games and Applications for Health (SeGAH)*, 2011 IEEE 1st International Conference on, 2011. IEEE, 1-4.
- ACT3D. 2015. *Quest3D* [Online]. ACT3D. Available: <http://www.quest3d.com/>.
- ACTIVISION. 2010. *Call of Duty* [Online]. Available: <https://www.activision.com/games/call-of-duty/call-of-duty-black-ops>.
- ADAMS, R., FINN, P., MOES, E., FLANNERY, K. & RIZZO, A. 2009. Distractibility in Attention/Deficit/ Hyperactivity Disorder (ADHD): The Virtual Reality Classroom. *Child Neuropsychology* 15.
- ADOBE. 2015. *Photoshop* [Online]. Adobe. Available: http://www.adobe.com/uk/products/photoshop.html?sdid=KKQPG&kw=semgeneri c&s_kwcid=AL!3085!3!69045922490!e!!g!!adobe%20photoshop&ef_id=VlsfoQ AAAfpcXvga:20151213133556:s.
- ALTSCHULER, E. L. 2008. Play with online virtual pets as a method to improve mirror neuron and real world functioning in autistic children. *Medical hypotheses*, 70, 748-749.
- AUTODESK. 2015a. *3DS MAX* [Online]. Autodesk. Available: <http://www.autodesk.co.uk/products/3ds-max/overview>.
- AUTODESK. 2015b. *Maya* [Online]. Autodesk. Available: <http://www.autodesk.co.uk/products/maya/overview>.
- AUTODESK. 2016. *MotionBuilder* [Online]. Available: <http://www.autodesk.com/products/motionbuilder/overview>.

- AYAZ, H., SHEWOKIS, P. A., BUNCE, S., IZZETOGLU, K., WILLEMS, B. & ONARAL, B. 2012. Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, 59, 36-47.
- BABU, S., SUMA, E., BARNES, T. & HODGES, L. F. Can immersive virtual humans teach social conversational protocols? Virtual Reality Conference, 2007. VR'07. IEEE, 2007. IEEE, 215-218.
- BAIENSON, J. N. & BEALL, A. C. 2006. Transformed social interaction: Exploring the digital plasticity of avatars. *Avatars at Work and Play*. Springer.
- BAIENSON, J. N., BEALL, A. C., LOOMIS, J., BLASCOVICH, J. & TURK, M. 2004. Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *PRESENCE: Teleoperators and Virtual Environments*, 13, 428-441.
- BAIENSON, J. N., BLASCOVICH, J., BEALL, A. C. & NOVECK, B. 2006a. Courtroom applications of virtual environments, immersive virtual environments, and collaborative virtual environments. *Law & Policy*, 28, 249-270.
- BAIENSON, J. N., YEE, N., MERGET, D. & SCHROEDER, R. 2006b. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators and Virtual Environments*, 15, 359-372.
- BANDAI. 2016. *tamagotchi-friends* [Online]. Bandai. Available: <https://www.bandai.co.uk/tamagotchi-friends/> [Accessed 02/05/2015 2016].
- BECERRA, E. P. & STUTTS, M. A. 2008. Ugly duckling by day, super model by night: The influence of body image on the use of virtual worlds. *Journal of Virtual Worlds Research*, 1, 1-19.

- BECK, S., KUNERT, A., KULIK, A. & FROEHLICH, B. 2013. Immersive group-to-group telepresence. *Visualization and Computer Graphics, IEEE Transactions on*, 19, 616-625.
- BEHREND, T., TOADDY, S., THOMPSON, L. F. & SHAREK, D. J. 2012. The effects of avatar appearance on interviewer ratings in virtual employment interviews. *Computers in Human Behavior*, 28, 2128-2133.
- BEKELE, E., ZHENG, Z., SWANSON, A., CRITTENDON, J., WARREN, Z. & SARKAR, N. 2013. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. *Visualization and Computer Graphics, IEEE Transactions on*, 19, 711-720.
- BENTE, G., RÜGGENBERG, S., KRÄMER, N. C. & ESCHENBURG, F. 2008. Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations. *Human communication research*, 34, 287-318.
- BERNSTEIN, D. M. & LOFTUS, E. F. 2009. How to tell if a particular memory is true or false. *Perspectives on Psychological Science*, 4, 370-374.
- BETHESDA. 2011. *Skyrim* [Online]. Bethesda Game Studios. Available: <http://www.elderscrolls.com/>.
- BIOCCA, F., BURGOON, J., HARMS, C. & STONER, M. 2001. Criteria and scope conditions for a theory and measure of social presence. *Presence: Teleoperators and virtual environments*.
- BLENDER. 2015. *Blender* [Online]. Available: <https://www.blender.org/>.
- BRENTON, H., GILLIES, M., BALLIN, D. & CHATTING, D. The uncanny valley: does it exist. Proceedings of conference of human computer interaction, workshop on human animated character interaction, 2005. Citeseer.

- BRENTON, H., GILLIES, M., SURMAN, D. & SLOAN, H. 2007. Uncanny Valley: Realism In Visualisation Character Design. *SwanQuake: The User Manual*, 93-100.
- BRYANT, D. 2005. The Uncanny Valley. Why are monster-movie zombies so horrifying and talking animals so fascinating. *Electronic resource available online: <http://www.arclight.net/~pdb/nonfiction/uncanny-valley.html>.*
- CALL, J. 2001. Chimpanzee social cognition. *Trends in cognitive sciences*, 5, 388-393.
- CANSIZ, Y. & TOKEL, S. T. Effects of Way Finding Affordances on Usability of Virtual World Environments in Terms of Users' Satisfaction, Performance, and Mental Workload: Examination by Eye-Tracking and fNIR. World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2012. 1073-1079.
- CASE, L., YANAGI, A., LOESER, E. & FULTZ, A. 2015. Human-animal relationships: The use of species-typical food calls and chimpanzee (*Pan troglodytes*) names: Welfare-oriented tools to manage sanctuary chimpanzees. *Animal Behavior and Cognition*, 2, 254-266.
- CF JR BOND, B. D. 2006. Accuracy of deception judgments. *Pers Soc Psychol Rev*, 10, 214-34.
- CHEN, Z. H., LIAO, C., CHIEN, T. C. & CHAN, T. W. 2011. Animal companions: Fostering children's effort-making by nurturing virtual pets. *British Journal of Educational Technology*, 42, 166-180.
- CHRISTIE. 2015. *MicroTiles* [Online]. Available: <https://www.christiedigital.com/en-us/business/products/microtiles>.
- COBI. 2015. *COBI Manual* [Online]. Available: <https://www.biopac.com/Manuals/fNIR%20COBI%20Manual.pdf>.
- COE, J., SCOTT, D. & LUKAS, K. 2009. Technical Report – Facility Design for Bachelor Gorilla Groups. *Zoo Biology* 28, 144-162.

- CONN, C., LANIER, J., MINSKY, M., FISHER, S. & DRUIN, A. Virtual environments and interactivity: Windows to the future. *ACM SIGGRAPH Computer Graphics*, 1989. ACM, 7-18.
- CROWELL, A. 1997. Gorillas in Bits. Available: <http://gtresearchnews.gatech.edu/reshor/rh-sf96/gorillas.htm>.
- CRUZ-NEIRA, C., SANDIN, D. J., DEFANTI, T. A., KENYON, R. V. & HART, J. C. 1992. The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, 35, 64-72.
- DALGARNO, B. & LEE, M. J. 2010. What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41, 10-32.
- DAUGHERTY, B., BABU, S., CUTLER, B. & HODGES, L. Officer Garcia: a virtual human for mediating eyewitness identification. *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, 2007. ACM, 117-120.
- DAUGHERTY, B., BABU, S., VAN WALLENDIAEL, L., CUTLER, B. & HODGES, L. F. 2008. Virtual human versus human administration of photographic lineups. *IEEE computer graphics and applications*, 65-75.
- DAVIES, G. 2000. 12 Virtual animals in electronic zoos. *Animal spaces, beastly places: New geographies of human-animal relations*, 10, 243.
- 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.

- DEAN, E., COOK, S., KEATING, M. & MURPHY, J. 2009. Does this avatar make me look fat? Obesity and interviewing in Second Life. *Journal For Virtual Worlds Research*, 2.
- DERLAGA, V. J. & BERG, J. H. 2013. *Self-disclosure: Theory, research, and therapy*, Springer Science & Business Media.
- DIETRICH, D. R. 2013. Avatars of Whiteness: Racial expression in video game characters. *Sociological Inquiry*, 83, 82-105.
- DILLAHUNT, T., BECKER, G., MANKOFF, J. & KRAUT, R. Motivating environmentally sustainable behavior changes with a virtual polar bear. Pervasive 2008 Workshop Proceedings, 2008. 58-62.
- DUCKWORTH, T. & ROBERTS, D. J. 2012. 3DRecon A utility for 3D reconstruction from video. *short paper at the Joint Virtual Reality Conference (ICAT, EEG, Eurographics)*. Madrid.
- DUCKWORTH, T. & ROBERTS, D. J. 2014. Parallel processing for real-time 3D reconstruction from video streams. *Journal of Real-Time Image Processing*, 9, 427-445.
- DWYER JR, F. 1968. Effect of varying amount of realistic detail in visual illustrations designed to complement programmed instruction. *Perceptual and Motor Skills*., 27, 351-354.
- DYCK, M., WINBECK, M., LEIBERG, S., CHEN, Y., GUR, R. C. & MATHIAK, K. 2008. Recognition profile of emotions in natural and virtual faces. *PLoS One*, 3, e3628.
- EKMAN, P. 1985. *Telling lies: Clues to deceit in the marketplace, marriage, and politics*. New York: Norton.
- EKMAN, P. 1993. Facial expression and emotion. *American psychologist*, 48, 384.

- EKMAN, P. 2003. Darwin, deception, and facial expression. *Annals of the New York Academy of Sciences*, 1000, 205-221.
- EKMAN, P. & FRIESEN, W. V. 1969. Nonverbal leakage and clues to deception. *Psychiatry*, 32, 88-106.
- EKMAN, P. & FRIESEN, W. V. 1978. *Facial Action Coding System: Investigator's Guide*, Consulting Psychologists Press.
- EKMAN, P. & MATSUMOTO, D. 2008. Facial expression analysis. *Scholarpedia*, 3, 4237.
- EKMAN, P. & YAMEY, G. 2004. Emotions revealed: recognising facial expressions. *careerfocus*, 328, 75-76.
- ENDRES, J. & LAIDLAW, A. 2009. Micro-expression recognition training in medical students: a pilot study. *BMC medical education*, 9, 47.
- EPICGAMES. *Unreal Engine 4* [Online]. Available: <https://www.unrealengine.com/what-is-unreal-engine-4>.
- EPRIME. 2015. *EPrime* [Online]. Available: <http://www.pstnet.com/eprime.cfm>.
- FACESHIFT. Available: <http://www.faceshift.com/>.
- FACESHIFT. 2015. *Markerless facial motion tracking system [Online]* [Online]. Available: <http://www.faceshift.com/> [Accessed 01/09/15].
- FACEWARE. 2015. *Faceware for UE4* [Online]. Available: <http://www.develop-online.net/tools-and-tech/key-release-faceware-live-for-ue4/0211750>.
- FAIRCHILD, A. J., CAMPION, S. P., GARCIA, A., WOLFF, R., FERNANDO, T. & ROBERTS, D. J. 2015. A Mixed Reality Telepresence System for Collaborative Space Operation.
- FALLOON, G. 2010. Using avatars and virtual environments in learning: What do they have to offer? . *British Journal of Educational Technology*, 41, 108-122.

- FELDMAN, R. S., JENKINS, L. & POPOOLA, O. 1979. Detection of deception in adults and children via facial expressions. *Child development*, 350-355.
- FERRARI, M. & QUARESIMA, V. 2012. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage*, 63, 921-935.
- FNIRDEVICES. 2016. *fNIR Imager 1000* [Online]. Available: <http://fnirdevices.com/>.
- FOGG, B. J. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002, 5.
- FOX, J. & BAILENSEN, J. 2010. The use of doppelgängers to promote health behavior change. *CyberTherapy & Rehabil*, 3, 16-17.
- FOX, J., BAILENSEN, J. N. & TRICASE, L. 2013. The embodiment of sexualized virtual selves: The Proteus effect and experiences of self-objectification via avatars. *Computers in Human Behavior*, 29, 930-938.
- FRANCO, J. & BOYER, E. 2009. Efficient polyhedral modeling from silhouettes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31, 414-427.
- FRANK, G., GUINN, C., HUBAL, R., POPE, P., STANFORD, M. & LAMM-WEISEL, D. Just-talk: An application of responsive virtual human technology. Proceedings of the Interservice/Industry Training, Simulation and Education Conference, 2002. 773-779.
- FRANK, M. G. & FEELEY, T. H. 2003. To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research*, 31, 58-75.
- FUCHS, H., BISHOP, G., ARTHUR, K., MCMILLAN, L., BAJCSY, R., LEE, S., FARID, H. & KANADE, T. Virtual space teleconferencing using a sea of cameras. Proc. First International Conference on Medical Robotics and Computer Assisted Surgery, 1994.

- GARAU, M., SLATER, M., VINAYAGAMOORTHY, V., BROGNI, A., STEED, A. & SASSE, M. A. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. Proceedings of the SIGCHI conference on Human factors in computing systems, 2003. ACM, 529-536.
- GILLIES, M. & SLATER, M. 2005. Non-verbal communication for correlational characters.
- GOOGLE. 2015. *Google Glass* [Online]. Available: <https://www.google.co.uk/intl/en/glass/start/>.
- GOOSSENS, B., SETCHELL, J., TCHIDONGO, E., DILAMBAKA, E., VIDAL, C., ANCRENAZ, M. & JAMART, A. 2005. Survival, interactions with conspecifics and reproduction in 37 chimpanzees released into the wild. *Biological conservation*, 123, 461-475.
- GRAAFLAND, M., SCHRAAGEN, J. & SCHIJVEN, M. P. 2012. Systematic review of serious games for medical education and surgical skills training. *British journal of surgery*, 99, 1322-1330.
- GROOM, V., NASS, C., CHEN, T., NIELSEN, A., SCARBOROUGH, J. K. & ROBLES, E. 2009. Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human-Computer Studies*, 67, 842-849.
- GUADAGNO, R. E., SWINTH, K. R. & BLASCOVICH, J. 2011. Social evaluations of embodied agents and avatars. *Computers in Human Behavior*, 27, 2380-2385.
- GUTIERREZ, J. 2011. *7 Films That Got Stuck In The Uncanny Valley* [Online]. Available: http://www.therobotsvoice.com/2011/03/7_films_that_got_stuck_in_the_uncanny_valley.php 2016].

- HALLORAN, J., FITZPATRICK, G., ROGERS, Y. & MARSHALL, P. Does it matter if you don't know who's talking?: multiplayer gaming with voiceover IP. CHI'04 extended abstracts on Human factors in computing systems, 2004. ACM, 1215-1218.
- HALLORAN, J., ROGERS, Y. & FITZPATRICK, G. From text to talk: multiplayer games and voiceover IP. DIGRA Conf., 2003.
- HARE, B., CALL, J., AGNETTA, B. & TOMASELLO, M. 2000. Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59, 771-785.
- HAY, K., CROZIER, J., BARNETT, M., ALLISON, D., BASHAW, M., HOOS, B. & PERKINS, L. 2000. Virtual Gorilla Modelling Project: Middle School Students Constructing Virtual Models for Learning. *Fourth International Conference of Learning Sciences*. .
- HELENO, P. & DOS SANTOS, M. P. 1998. Artificial animals in virtual ecosystems. *Computer networks and ISDN systems*, 30, 1923-1932.
- HEW, K. F. & CHEUNG, W. S. 2010. Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British journal of educational technology*, 41, 33-55.
- HINDE, R. 1972. *Non-Verbal Communication*, Cambridge University Press.
- HORN, D. B. Is seeing believing?: detecting deception in technologically mediated communication. CHI'01 Extended Abstracts on Human Factors in Computing Systems, 2001. ACM, 297-298.
- HTC. 2015. *HTC VIVE* [Online]. Available: <http://www.htevr.com/>.
- HU, F., WANG, Y., CAO, X., HAO, Q. & BROWN, D. 2013. Using Wiimote and Kinect for Cognitive Rehabilitation: Toward an Intelligent Sensor/Image Signal Processing. *Cyber-Physical Systems: Integrated Computing and Engineering Design*, 321.

- IERONUTTI, L. & CHITTARO, L. 2007. Employing Virtual Humans for education and training in X3D/VRML Worlds. *Computers and Education* 49, 93-109.
- IMMERSIVEDISPLAY. 2015. *Dome Displays* [Online]. Available: <http://immersivedisplay.co.uk/ImmersiveDisplayGroup/index.php/domes/>.
- IRENAEUS, J. A., BOEKHORST, T. & HOGEWEG, P. 1994. Self-structuring in artifactle “chimps” offers new hypothesis for male grouping in chimpanzees. *Behaviour*, 130, 3-4.
- JAIN, S., TAMERSOY, B., ZHANG, Y., AGGARWAL, J. & ORVALHO, V. An interactive game for teaching facial expressions to children with Autism Spectrum Disorders. *Communications Control and Signal Processing (ISCCSP), 2012 5th International Symposium On, 2012. IEEE, 1-4.*
- JENVOID, M. 2008. Chimpanzee (Pan Troglodytes) Response to Caregivers use of Chimpanzee Behaviour. *Zoo Biology* 27, 345-359.
- KAPLAN, F. Free creatures: The role of uselessness in the design of artificial pets. *Proceedings of the 1st edutainment robotics workshop, 2000. Citeseer, 45-47.*
- KETELHUT, D. J. & SCHIFTER, C. C. 2011. Teachers and game-based learning: Improving understanding of how to increase efficacy of adoption. *Computers & Education*, 56, 539-546.
- KHANNA, P., YU, I., MORTENSEN, J. & SLATER, M. Presence in response to dynamic visual realism: a preliminary report of an experiment study. *Proceedings of the ACM symposium on Virtual reality software and technology, 2006. ACM, 364-367.*
- KINECT, M. *Microsoft Kinect* [Online]. Available: <http://www.microsoft.com/en-us/kinectforwindows/>.
- KITAGAWA, M. & WINDSOR, B. 2008. *MoCap for Artists Workflow and Techniques for Motion Capture*, Focal Press.

- KUSAHARA, M. 2001. The art of creating subjective reality: an analysis of Japanese digital pets. *Leonardo*, 34, 299-302.
- LANTZ, E. The future of virtual reality: head mounted displays versus spatially immersive displays (panel). Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996. ACM, 485-486.
- The Lawnmower Man*, 1992. Directed by LEONARD, B.
- LIFE, S. 2015. *The Open University Case Study* [Online]. Available: http://wiki.secondlife.com/wiki/Case_Study:_The_Open_University.
- LIM, S. & REEVES, B. 2010. Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. *International Journal of Human-Computer Studies*, 68, 57-68.
- LLOBERA, J., SPANLANG, B., RUFFINI, G. & SLATER, M. 2010. Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception (TAP)*, 8, 3.
- MAIMONE, A. & FUCHS, H. Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras. Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on, 2011. IEEE, 137-146.
- MARTI, P., POLLINI, A., RULLO, A. & SHIBATA, T. Engaging with artificial pets. Proceedings of the 2005 annual conference on European association of cognitive ergonomics, 2005. University of Athens, 99-106.
- MATSUMOTO, D. & HWANG, H. S. 2011. Reading facial expressions of emotion. *Psychological Science Agenda*, 5, 10-18.
- MCDONNELL, R., BREIDT, M. & BULTHOFF, H. Render me Real? Investigating the Effect of Render Style on the Perception of Animated Virtual Humans. *ACM Transactions on Graphics*, 2012. 1-11.

- MCKENZIE, F. R., SCERBO, M., CATANZARO, J. & PHILLIPS, M. 2003. Nonverbal indicators of malicious intent: affective components for interrogative virtual reality training. *International journal of human-computer studies*, 59, 237-244.
- MELFI, V. 2009. There are big gaps in our knowledge, and thus approach, to zoo animal welfare: a case for evidence-based zoo animal management. *Zoo biology*, 28, 574-588.
- MICROSOFT. *Hololens* [Online]. Microsoft. Available: <https://www.microsoft.com/microsoft-hololens/en-us>.
- MIKROPOULOS, T. A. & NATSIS, A. 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, 56, 769-780.
- MILGRAM, P. & KISHINO, F. 1994. A Taxonomy of Mixed Reality Visual-Displays. *Ice Transactions on Information and Systems*, E77d, 1321-1329.
- MILLAR, M. G. & MILLAR, K. U. 1997. The effects of cognitive capacity and suspicion on truth bias. *Communication Research*, 24, 556-570.
- MILLINGTON, I. & FUNGE, J. 2009. *Artifical Intelligence for Games*
- MITANI, J. C. 2009. Male chimpanzees form enduring and equitable social bonds. *Animal Behaviour*, 77, 633-640.
- MITANI, J. C. & WATTS, D. P. 2001. Why do chimpanzees hunt and share meat? *Animal Behaviour*, 61, 915-924.
- MITANI, J. C., WATTS, D. P., PEPPER, J. W. & MERRIWETHER, D. A. 2002. Demographic and social constraints on male chimpanzee behaviour. *Animal Behaviour*, 64, 727-737.
- MOORE, C., DUCKWORTH, T., ASPIN, R. & ROBERTS, D. Synchronization of images from multiple cameras to reconstruct a moving human. *Distributed Simulation and*

- Real Time Applications (DS-RT), 2010 IEEE/ACM 14th International Symposium on, 2010. IEEE, 53-60.
- MORGAN, P., CLEAVE-HOGG, D., MCLLROY, J. & DEVITT, J. 2002. Simulation Technology a comparison of experiential and visual learning for undergraduate medical students. *Anesthesiology*, 96, 10-16.
- MURRAY, C. M., MANE, S. V. & PUSEY, A. E. 2007. Dominance rank influences female space use in wild chimpanzees, *Pan troglodytes*: towards an ideal despotic distribution. *Animal Behaviour*, 74, 1795-1804.
- MURRAY, N., ROBERTS, D., STEED, A., SHARKEY, P., DICKERSON, P., RAE, J. & WOLFF, R. 2009. Eye gaze in virtual environments: evaluating the need and initial work on implementation. *Concurrency and Computation: Practice and Experience*, 21, 1437-1449.
- NASOZ, F. & LISETTI, C. L. 2006. MAUI avatars: Mirroring the user's sensed emotions via expressive multi-ethnic facial avatars. *Journal of Visual Languages & Computing*, 17, 430-444.
- NINTENDO. *Nintendogs* [Online]. Available: <http://nintendogspluscats.nintendo.com/>.
- NINTENDO. 2007. *Wii Fit Plus* [Online]. Available: <http://wiifit.com/training/yoga.html>.
- NJENGA, J. K. & FOURIE, L. C. H. 2010. The myths about e-learning in higher education. *British Journal of Educational Technology*, 41, 199-212.
- NOWAK, K. L. & BIOCCA, F. 2003. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence*, 12, 481-494.
- O'HARE, J. 2015. *Octave - technical information, University of Salford [Online]* [Online]. Available: <http://www.salford.ac.uk/computing-science-engineering/facilities/octave-technical-information> [Accessed 01/09/15].

- OCULUS. *Oculus Rift* [Online]. Available: <https://www.oculus.com/en-us/> [Accessed 29/11/2015 2015].
- ORVALHO, V., MIRANDA, J. & SOUSA, A. S. 2010. What a feeling: Learning facial expressions and emotions. *Revista PRISMA. COM*.
- OXFORDDICTIONARIES.COM. *Virtual Reality Definition* [Online]. Oxford Dictionary. Available: <http://www.oxforddictionaries.com/definition/english/virtual-reality>.
- PAN, X., GILLIES, M. & SLATER, M. The impact of avatar blushing on the duration of interaction between a real and virtual person. *Presence 2008: The 11th Annual International Workshop on Presence, 2008*. Citeseer, 100-106.
- PAN, X. & SLATER, M. A preliminary study of shy males interacting with a virtual female. *Presence: The 10th Annual International Workshop on Presence, 2007*.
- PAN, Z., CHEOK, A., YANG, H., ZHU, J. & SHI, J. 2006a. Virtual Reality and Mixed Reality for Virtual Learning Environments. *Computer and Graphics* 30, 20-28.
- PAN, Z., CHEOK, A. D., YANG, H., ZHU, J. & SHI, J. 2006b. Virtual reality and mixed reality for virtual learning environments. *Computers & graphics*, 30, 20-28.
- PARR, L. A. & WALLER, B. M. 2006. Understanding chimpanzee facial expression: insights into the evolution of communication. *Social Cognitive and Affective Neuroscience*, 1, 221-228.
- PARR, L. A., WALLER, B. M., VICK, S. J. & BARD, K. A. 2007. Classifying chimpanzee facial expressions using muscle action. *Emotion*, 7, 172.
- PAUSCH, R., SNODDY, J., TAYLOR, R., WATSON, S. & HASELTINE, E. Disney's Aladdin: first steps toward storytelling in virtual reality. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996*. ACM, 193-203.

- PEÑA, J. & HANCOCK, J. T. 2006. An analysis of socioemotional and task communication in online multiplayer video games. *Communication Research*, 33, 92-109.
- PEÑA, J., MCGLONE, M. S. & SANCHEZ, J. 2012. The cowl makes the monk: How avatar appearance and role labels affect cognition in virtual worlds. *Journal For Virtual Worlds Research*, 5.
- PHILIPP, M. C., STORRS, K. R. & VANMAN, E. J. 2012. Sociality of facial expressions in immersive virtual environments: A facial EMG study. *Biological psychology*, 91, 17-21.
- PICAULT, S. & COLLINOT, A. Designing social cognition models for multi-agent systems through simulating primate societies. Multi Agent Systems, 1998. Proceedings. International Conference on, 1998. IEEE, 238-245.
- PORTER, S. & TEN BRINKE, L. 2008. Reading between the lies identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, 19, 508-514.
- POZNANSKI*, M. & THAGARD, P. 2005. Changing personalities: towards realistic virtual characters. *Journal of Experimental & Theoretical Artificial Intelligence*, 17, 221-241.
- PREMACK, D. & PREMACK, A. J. 1994. Levels of causal understanding in chimpanzees and children. *Cognition*, 50, 347-362.
- PUGA-GONZALEZ, I., HILDENBRANDT, H. & HERNELRIJK, C. 2009. Emergent Patterns of Social Affiliation in Primates, A Model. *PLos Computational Biology*, 5.
- QUALISYS. 2015. *Qualisys Motion Tracking* [Online]. Available: <http://www.qualisys.com/>.

- RAGAN, E. D., HUBER, K. J., LAHA, B. & BOWMAN, D. The effects of navigational control and environmental detail on learning in 3D virtual environments. *Virtual Reality Short Papers and Posters (VRW)*, 2012 IEEE, 2012. IEEE, 11-14.
- RASKAR, R., WELCH, G., CUTTS, M., LAKE, A., STESIN, L. & FUCHS, H. The office of the future: A unified approach to image-based modeling and spatially immersive displays. *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998. ACM, 179-188.
- RIZZO, A. & KIM, G. J. 2005. A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. *PRESENCE*, 14, 119-146.
- ROBERTS, D., RAE, J., DUCKWORTH, T., MOORE, C. & ASPIN, R. 2013a. Estimating the Gaze of a Virtuality Human. *IEEE Transactions on Visualization and Computer Graphics*, 19, 681-690.
- ROBERTS, D., WOLFF, R., OTTO, O. & STEED, A. 2003. Constructing a Gazebo: supporting teamwork in a tightly coupled, distributed task in virtual reality. *Presence: Teleoperators and Virtual Environments*, 12, 644-657.
- ROBERTS, D. J., FAIRCHILD, A. J., CAMPION, S. P., O'HARE, J., MOORE, C. M., ASPIN, R., DUCKWORTH, T., GASPARELLO, P. & TECCHIA, F. 2015a. withyou—An Experimental End-to-End Telepresence System Using Video-Based Reconstruction. *Selected Topics in Signal Processing, IEEE Journal of*, 9, 562-574.
- ROBERTS, D. J., GARCIA, A. S., DODIYA, J., WOLFF, R., FAIRCHILD, A. J. & FERNANDO, T. Collaborative telepresence workspaces for space operation and science. *Virtual Reality (VR)*, 2015 IEEE, 2015b. IEEE, 275-276.
- ROBERTS, D. J., RAE, J., DUCKWORTH, T. W., MOORE, C. M. & ASPIN, R. 2013b. Estimating the gaze of a virtuality human. *Visualization and Computer Graphics, IEEE Transactions on*, 19, 681-690.

- ROSENBERG, R. S., BAUGHMAN, S. L. & BAIENSON, J. N. 2013. Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PloS one*, 8, e55003.
- ROSS, S. & GILLESPIE, K. 2009. Research Article – Influences on Visitor Behaviour at a Modern Immersive Zoo Exhibit. *Zoo Biology* 28, 462-472.
- SAMSUNG. 2015. *Curved TV* [Online]. Available: <http://www.samsung.com/uk/consumer/tv-audio-video/televisions/suhd-nano-crystal-tvs/UE65JS9000TXXU>.
- SERAGLIA, B., GAMBERINI, L., PRIFTIS, K., SCATTURIN, P., MARTINELLI, M. & CUTINI, S. 2011. An exploratory fNIRS study with immersive virtual reality: a new method for technical implementation. *Frontiers in human neuroscience*, 5.
- SHOGAKUKAN. 2008. *Professor Kageyama's Maths Training* [Online]. Available: <https://www.nintendo.co.uk/Games/Nintendo-DS/Professor-Kageyama-s-Maths-Training-272530.html#Home>.
- SIMS, E. M. 2007. Reusable, lifelike virtual humans for mentoring and role-playing. *Computers & Education*, 49, 75-92.
- SLATER, M. 2003. A note on presence terminology. *Presence connect*, 3, 1-5.
- 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.
- SLATER, M., BROGNI, A. & STEED, A. Physiological responses to breaks in presence: A pilot study. *Presence 2003: The 6th Annual International Workshop on Presence*, 2003. Citeseer.

- SLATER, M., KHANNA, P., MORTENSEN, J. & YU, I. 2009. Visual realism enhances realistic response in an immersive virtual environment. *Computer Graphics and Applications, IEEE*, 29, 76-84.
- SLATER, M., ROVIRA, A., SOUTHERN, R., SWAPP, D., ZHANG, J. J., CAMPBELL, C. & LEVINE, M. 2013. Bystander responses to a violent incident in an immersive virtual environment. *PloS one*, 8, e52766.
- SLATER, M. & STEED, A. 2000. A virtual presence counter. *Presence*, 9, 413-434.
- SMITH, D., MCLAUGHLIN, T. & BROWN, I. 2012a. 3-D Computer Animation Vs Live-Action Video: Differences in Viewers Response to Instructional Vignettes. *Contemporary Issues in Technology and Teacher Education.*, 12, 41-54.
- SMITH, D., MCLAUGHLIN, T. & BROWN, I. 2012b. 3-D Computer Animation vs. Live-Action Video: Differences in Viewers' Response to Instructional Vignettes. *Contemporary Issues in Technology and Teacher Education*, 12, 41-54.
- SONY. 2015. *Sony Morpheus* [Online]. Sony. Available: <https://www.playstation.com/en-us/explore/playstation-vr/>.
- SPELLMAN, B. A. & TENNEY, E. R. 2010. Credible testimony in and out of court. *Psychonomic bulletin & review*, 17, 168-173.
- SSAT. 2010. *VIBE* [Online]. Available: <http://www.ssatvibe.org.uk/>.
- STEPTOE, W., STEED, A., ROVIRA, A. & RAE, J. Lie tracking: social presence, truth and deception in avatar-mediated telecommunication. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2010. ACM, 1039-1048.
- STEUER, J. 1992. Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of communication*, 42, 73-93.

- STRICKLIN, W., ZHOU, J. & GONYOU, H. 1995. Selfish animats and robot ethology: using artificial animals to investigate social and spatial behavior. *Applied Animal Behaviour Science*, 44, 187-203.
- SUSI, T., JOHANNESSON, M. & BACKLUND, P. 2007. Serious games: An overview.
- THOMPSON, C. 2004. The undead zone: Why realistic graphics make humans look creepy. *SLATE* (June 9, 2004).
- TINWELL, A., GRIMSHAW, M., NABI, D. A. & WILLIAMS, A. 2011. Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, 27, 741-749.
- TOMASELLO, M., CALL, J. & HARE, B. 2003a. Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends in cognitive sciences*, 7, 153-156.
- TOMASELLO, M., CALL, J. & HARE, B. 2003b. Chimpanzees versus humans: it's not that simple. *Trends in Cognitive Sciences*, 7, 239-240.
- TOMASELLO, M., HARE, B. & AGNETTA, B. 1999. Chimpanzees, Pan troglodytes, follow gaze direction geometrically. *Animal Behaviour*, 58, 769-777.
- TOMASELLO, M., HARE, B. & FOGLEMAN, T. 2001. The ontogeny of gaze following in chimpanzees, Pan troglodytes, and rhesus macaques, Macaca mulatta. *Animal Behaviour*, 61, 335-343.
- TOMLINSON, B. & BLUMBERG, B. 2001. Social behaviour, emotion and learning in a pack of virtual wolves.
- TROMP, J., BULLOCK, A., STEED, A., SADAGIC, A., SLATER, M. & FRÉCON, E. 1998. Small group behaviour experiments in the coven project. *Computer Graphics and Applications, IEEE*, 18, 53-63.

- UBISOFT. 2008. *Gourmet Chef* [Online]. Available: <http://www.nintendo.com/games/detail/abrTWecMFBvDfQhQ47EV03htfJbRnhLC>.
- UNITY3D. 2015. *Unity3D* [Online]. Available: <http://unity3d.com/>.
- VICK, S.-J., WALLER, B. M., PARR, L. A., PASQUALINI, M. C. S. & BARD, K. A. 2007. A cross-species comparison of facial morphology and movement in humans and chimpanzees using the facial action coding system (FACS). *Journal of Nonverbal Behavior*, 31, 1-20.
- VICON. Available: <http://www.vicon.com/Press/Details/10>.
- VICON. 2015. *Vicon optical motion tracking system [Online]* [Online]. Available: <http://www.vicon.com/> [Accessed 01/09/15].
- VINAYAGAMOORTHY, V., GARAU, M., STEED, A. & SLATER, M. An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience. *Computer Graphics Forum*, 2004. Wiley Online Library, 1-11.
- VINAYAGAMOORTHY, V., GILLIES, M., STEED, A., TANGUY, E., PAN, X., LOSCOS, C. & SLATER, M. 2006. Building expression into virtual characters.
- VIRTUIX. 2015. *OMNI Treadmill* [Online]. Available: <http://www.virtuix.com/>.
- VISSCHER, N. C., SNIDER, R. & VANDER STOEP, G. 2009. Comparative analysis of knowledge gain between interpretive and fact-only presentations at an animal training session: an exploratory study. *Zoo biology*, 28, 488-495.
- VRIJ, A. & MANN, S. 2001. Who killed my relative? Police officers' ability to detect real-life high-stake lies. *Psychology, Crime and Law*, 7, 119-132.
- WAZA 2005. *Building a Future for Wildlife - The World Zoo and Aquarium Conservation Strategy*.

- WESCHLER, L. 2002. Why is this man smiling. *Digital animators are closing in on the complex system that makes a face come alive. Wired*, 10.
- WORLDS, A. 2015. *River City Project* [Online]. Available: <http://rivercity.activeworlds.com/>.
- XVR. 2015. *XVR virtual reality training software for training and security* [Online]. Available: <http://www.xvrsim.com/>.
- YALON-CHAMOVITZ, S. & WEISS, P. L. T. 2008. Virtual reality as a leisure activity for young adults with physical and intellectual disabilities. *Research in Developmental Disabilities*, 29, 273-287.
- YAP, M. H., UGAIL, H. & ZWIGGELAAR, R. 2012. Facial Analysis for Real-Time Application: A Review in Visual Cues Detection Techniques. *Journal of Communication and Computer*, 9, 1231-1241.
- YEE, N., BAIENSON, J. N., URBANEK, M., CHANG, F. & MERGET, D. 2007. The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior*, 10, 115-121.

Appendices

Appendix 1 - Information sheet and consent form for experiment 1

Appendix 2 – Chimpanzee Behaviours

Appendix 2 - Questionnaire for Experiment 1

Appendix 3 - Ethical Approval for Experiment 1

Appendix 4 - Recruitment email for Experiment 2

Appendix 5 - Information sheet and consent form for experiment 2 confederate

Appendix 6 - Information sheet and consent form for experiment 2 participant

Appendix 7 - Ethical Approval for Experiment 2

Appendix 1: Information sheet and consent form for experiment 1

University of
Salford
MANCHESTER

**The Centre for Virtual Environments and
Future Media.
Simon Campion BSc MSc**

Contact details: s.p.campion@salford.ac.uk
Web address: http://www.cve.salford.ac.uk/page/Simon_Campion

Information sheet for Participants

Participant Number []

Thank you for volunteering to participate in this study. Please read through this information and feel free to ask the experimenters any questions. This study will take place using projection equipment, and a range of display devices.

Procedures

- You will be asked to sign a consent form
- You will be asked to watch two short videos
- You will be interviewed by 2 interviewers
- Afterwards there will be a brief de-brief with the experimenter

The whole study is estimated to take about 20 minutes to complete however you should not feel under any pressure to rush. Take as much time as you need.

Please note

- No identifying information about participants will be published in any form
- You are free to withdraw from the study at any time and without giving reasons for withdrawing.

A note about virtual reality equipment

When using projection systems, people sometimes experience a degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please say so and we will stop.

With any type of video equipment there is a possibility that an epileptic episode may be generated. This, for example, has been reported for computer and video games. For this reason we regret that we are unable to accept volunteers who are known to have suffered from epilepsy.

Please turn over for the experiment consent form.....

Consent Form

Project: PhD Project – Socially Situated Simian Simulation

To be completed by volunteers: We would like you to read the following questions carefully.

Have you read the information sheet about this study?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all your questions?	YES/NO
Have you received enough information about this study?	YES/NO
Which investigator have you spoken to about this study? _____	
Do you understand that you are free to withdraw from this study?	
• At any time	YES/NO
• Without giving a reason for withdrawing	YES/NO
Do you agree to take part in this study?	YES/NO
Do you agree to be video taped?	YES/NO
Do you agree to be audio taped?	YES/NO
Do you agree to the data being stored and used in the researcher's ongoing research?	YES/NO
Do you agree to the data being stored indefinitely?	YES/NO
Do you agree to let us make transcripts of your interactions and present them in printed publications?	YES/NO
Do you agree to let us make transcripts of your interactions and present them in presentations?	YES/NO
Do you agree to let us use audio fragments involving you in presentations?	YES/NO
Do you agree to let us use video frames or photo stills of you in printed publications?	YES/NO
Do you agree to let us use the video or photo stills of you in presentations?	YES/NO
Do you agree to the audio, video and other data being shared with academic collaborators on the project at University of Salford?	YES/NO
Do you agree to completing a personality questionnaire?	YES/NO

I certify that I do not have epilepsy.

Signed..... Date.....

Name in block letters.....

Investigator.....

In case you have any enquiries regarding this study in the future, please contact:

Simon Champion
The Centre for Virtual Environments and Future Media,
Media City UK,
Salford Quays,
University of Salford,
Salford, Greater Manchester,
+44 161 295 2931
Email: s.p.campion@salford.ac.uk

Appendix 2: Chimpanzee Behaviours

University of
Salford
MANCHESTER

**The Centre for Virtual Environments and
Future Media.
Simon Campion BSc MSc**

Contact details: s.p.campion@salford.ac.uk
Web address: http://www.cve.salford.ac.uk/page/Simon_Campion

Behaviours to try and identify during the study

Please read the following behavioural descriptions before taking part in the study.

Approach (APP)(Type C): Goodall (1989): "Individual moves toward another." Plooijs's approaching (APP). Van Hooff (1973) divided approach into "smooth" and "hesitant." Includes <walk on all fours>, <walk on one's feet>, and <run>. Cf. Leave. Category 1.

Bob (BOB)(Type A): Goodall (1989): "The body bobs up and down as elbows are flexed and straightened. Shown typically by adolescent males but also by adolescent females and juvenile and adult males when a high-ranking individual passes. Sometimes as the dominant recipient of the gesture moves on, the bobber backs away in front of the dominant animal. This may provoke an aggressive response. Bobbing is accompanied by pant-grunts which may become rather frenzied pant-screams." Plooijs's (1984) Bob (BOB). Van Hooff's (1973) Squat bob. Absent in bonobos. Category 3.

Eat leaves (EAL)(Type B): Put leaves into the mouth and ingest them. Chimpanzee may bring mouth to the leaves or bring cropped leaves into the mouth. Important cropping technique is <pull through>. Category 1.

Groom (GRM) (Type B): Goodall's (1989) Grooming behavior: "...use both hands, pushing the hair back with the thumb or index finger of one hand and holding it back while picking at the exposed skin with the nail of the thumb or index finger of the other. The chimpanzee can also use one hand, parting the hair in the same way and holding it back with the lower lip". Grooming may occur in bipedal, quadrupedal, sitting, or lying posture. Function of grooming includes appeasement, reassurance and reconciliation in addition to elimination of ectoparasites. See Fig.6 of Nishida (1983a) for photograph. Category 1.

Rest (RES) (Type B): Goodall's (1989) Rest: Remain immobile, <sit> or <lie>. Kano's (1998) Resting. Category 1.

Sit (SIT) (Type A): Plooijs's (1984) SIT. Goodall (1989): "Rump on ground, branch, etc., body more or less vertical." See Fig.14a, c, d, g of Goodall (1968) for illustration. Kano's (1998) Sit. Category 1.

Walk quadrupedal with knuckle (WQK) (Type A): Goodall's (1989) Knuckle walking, Plooijs's (1984) WAQ. Quadrupedal walk with knuckles on the ground or large horizontal bough. See Fig. 6 of Nishida (1994). Kano's (1998) Walk. Category 2.

Run quadrupedal (RUQ) (Type A): Goodall's (1989) Quadrupedal run. Category 2?

NOTE: All behavioural descriptions were taken from the Ethogram and Ethnography of Mahale Chimpanzees, Nishida et al. Sub-Department of Anthropology, Faculty of Science, Kyoto University, Kitashirakawa-Oiwakecho, Sakyo, Kyoto, Japan.

Appendix 2: Questionnaire for Experiment 1

University of
Salford
MANCHESTER

**The Centre for Virtual Environments and
Future Media.
Socially Situated Simian Simulation**

Participant number []

Questions about you:

Male [] Female []

Age: 20-30 30-40 40-50 50-60 60-70

Occupation: _____

Do you have any eyesight problems at the time of the experiment? (Corrected by Glasses or contacts worn during the experiment)

Do you have any injuries or disabilities that you feel may have hindered your performance in the experiment if so please give details?

How often do you watch film and TV?

Never [] Occasionally [] Frequently [] Excessively []

How often do you use computers?

Never [] Occasionally [] Frequently [] Excessively []

How often do you play computer games?

Never [] Occasionally [] Frequently [] Excessively []

How often do you watch wildlife documentaries?

Never [] Occasionally [] Frequently [] Excessively []

How often do you visit the zoo?

Never [] Occasionally [] Frequently [] Excessively []

How interested are you in animal behaviour?

No Interest [] Interested [] Very Interested []

Any other relevant information:

Questions about your experience with this experiment:

When watching the video which of the following behaviours from the behaviour sheet could you identify? (Please tick the appropriate boxes)

Walk Quadrupedal [] Rest [] Eat Leaves [] Grooming []
Approach [] Sit [] Run Quadrupedal [] Bob []

When using the virtual environment which of the following behaviours from the behaviour sheet could you identify? (Please tick the appropriate boxes)

Walk Quadrupedal [] Rest [] Eat Leaves [] Grooming []
Approach [] Sit [] Run Quadrupedal [] Bob []

To what extent would you agree with the following statements? Please circle the appropriate answer. To strongly agree select 7, to neither agree nor disagree select 4 and to strongly disagree select 1.

1. The use of the virtual environment evoked a strong sense of engagement:

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

2. If I needed to learn more about animal behaviour I would prefer to use a video rather than a virtual environment.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

3. When using the virtual environment I could identify different animal behaviours such as walking, grooming, feeding and resting etc.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

4. When watching the video I could identify different animal behaviours such as walking, grooming, feeding and resting etc.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

5. The visual fidelity of the virtual environment was sufficient to identify animal behaviours.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

6. Watching the video made me curious to learn more about the animals.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

7. The use of the virtual environment made the task more enjoyable than watching the video.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

8. When using the virtual environment I felt as though I was in the animal enclosure.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

9. If I needed to learn more about animal behaviour I would prefer to use a virtual environment rather than a video.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

10. The visual quality of the virtual environment was insufficient to identify animal behaviours.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

11. The use of the video invoked a strong sense of engagement:

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

12. Using the virtual environment made me curious to learn more about the animals.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

13. I would like to see virtual environment technology used at zoos as part of talks or demonstrations.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

14. The animals were animated to an insufficient standard to allow behaviours to be identified within the virtual environment.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

15. When watching the video I felt as though I was in the animal enclosure.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

16. I would like to see virtual environment technology used online as a learning resource for zoos or educational institutions.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

17. The use of the video made the task more enjoyable than using the virtual environment.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

18. The animals were animated to a sufficient standard to allow for behaviour identification within the virtual environment.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

19. I would like to see video technology used at zoos as part of talks or demonstrations.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

20. I would like to see video used online as a learning resource for zoos or educational institutions.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

Any additional comments:

Thank you for taking time to participate in this study.

Appendix 3: Ethical Approval for Experiment 1

Academic Audit and Governance Committee

College of Science and Technology Research Ethics Panel
(CST)

University of
Salford
MANCHESTER

To Simon Campion and Prof David Roberts
cc: Prof Sunil Vadera, Head of School of CSE
From Nathalie Audren Howarth, College Research Support Officer
Date 20 February 2013


MEMORANDUM

Subject: Approval of your Project by CST
Project Title: Socially Situated Simian Simulation
REP Reference: CST 12/51

Following your responses to the Panel's queries, based on the information you provided, I can confirm that they have no objections on ethical grounds to your project.

If there are any changes to the project and/or its methodology, please inform the Panel as soon as possible.

Regards,


Nathalie Audren Howarth
College Research Support Officer

For enquiries please contact:
College of Science and Technology
College Research Support Officer
The University of Salford
Maxwell building, (7th floor, room 721)
Telephone: 0161 295 5278
Email: n.audren@salford.ac.uk

Appendix 4: Recruitment email for Experiment 2

Dear,

I am conducting an experiment about detecting deception on (.....) at Allerton building and wondered whether you would have time to take part? The study will take approximately 30 minutes to complete and involves watching a series of movie clips whilst wearing a head band that measures cognitive load. I have time slots available in the at If you have any questions then please don't hesitate to contact me. I look forward to hearing from you.

Kind regards

Simon

Simon Champion B.Sc, M.Sc

VR Consultant | VETS3D

THINKlab, University of Salford, 7th Floor Maxwell Building, M5 4WT

t: +44 (0) 161 295 2931

s.p.champion@salford.ac.uk | www.salford.ac.uk

<http://www.youtube.com/user/VETS3D>

http://www.cve.salford.ac.uk/page/Simon_Campion



Appendix 5: Information sheet and consent form for experiment 2 confederate

University of
Salford
MANCHESTER

**The Centre for Virtual Environments and
Future Media.
Simon Campion BSc MSc**

Contact details: s.p.campion@salford.ac.uk
Web address: http://www.cve.salford.ac.uk/page/Simon_Campion

Information sheet for Participants

Participant Number []

Thank you for volunteering to participate as a confederate in this study. Please read through this information and feel free to ask the experimenters any questions. This study will take place using projection equipment, and a range of display devices.

Procedures

- You will be asked to sign a consent form
- You will be asked to wear equipment to capture Physiological data.
- You will be asked to answer a series of questions
- Afterwards there will be a brief de-brief with the experimenter

The whole study is estimated to take about 20 minutes to complete however you should not feel under any pressure to rush. Take as much time as you need.

Please note

- No identifying information about participants will be published in any form
- You are free to withdraw from the study at any time and without giving reasons for withdrawing.

A note about virtual reality equipment

When using projection systems, people sometimes experience a degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please say so and we will stop.

With any type of video equipment there is a possibility that an epileptic episode may be generated. This, for example, has been reported for computer and video games. For this reason we regret that we are unable to accept volunteers who are known to have suffered from epilepsy.

Please turn over for the experiment consent form.....

Consent Form

Project: PhD Project – Socially Situated Simian Simulation

To be completed by volunteers: We would like you to read the following questions carefully.

Have you read the information sheet about this study?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all your questions?	YES/NO
Have you received enough information about this study?	YES/NO
Which investigator have you spoken to about this study? _____	
Do you understand that you are free to withdraw from this study?	
• At any time	YES/NO
• Without giving a reason for withdrawing	YES/NO
Do you agree to take part in this study?	YES/NO
Do you agree to be video taped?	YES/NO
Do you agree to be audio taped?	YES/NO
Do you agree to your heart rate being recorded?	YES/NO
Do you agree to your cognitive load being recorded?	YES/NO
Do you agree to the data being stored and used in the researcher's ongoing research?	YES/NO
Do you agree to the data being stored indefinitely?	YES/NO
Do you agree to let us make transcripts of your interactions and present them in printed publications?	YES/NO
Do you agree to let us make transcripts of your interactions and present them in presentations?	YES/NO
Do you agree to let us use audio fragments involving you in presentations?	YES/NO
Do you agree to let us use video frames or photo stills of you in printed publications?	YES/NO
Do you agree to let us use the video or photo stills of you in presentations?	YES/NO
Do you agree to the audio, video and other data being shared with academic collaborators on the project at University of Salford?	YES/NO
Do you agree to completing a personality questionnaire?	YES/NO

I certify that I do not have epilepsy.

Signed..... Date.....

Name in block letters.....

Investigator.....

In case you have any enquiries regarding this study in the future, please contact:

Simon Champion
The Centre for Virtual Environments and Future Media,
THINKLab,
7th Floor Maxwell Building,
University of Salford,
Salford, Greater Manchester, M5 4WT
+44 161 295 2931
Email: s.p.champion@salford.ac.uk

Appendix 6: Information sheet and consent form for experiment 2 participant

University of
Salford
MANCHESTER

**The Centre for Virtual Environments and
Future Media.
Simon Campion BSc MSc**

Contact details: s.p.campion@salford.ac.uk
Web address: http://www.cve.salford.ac.uk/page/Simon_Campion

Information sheet for Participants

Participant Number []

Thank you for volunteering to participate in this study. Please read through this information and feel free to ask the experimenters any questions. This study will take place using projection equipment, and a range of display devices.

Procedures

- You will be asked to sign a consent form
- You will be asked to wear equipment to capture Physiological data.
- You will be asked to watch short videos
- You will be interviewed by 2 interviewers
- Afterwards there will be a brief de-brief with the experimenter

The whole study is estimated to take about 20 minutes to complete however you should not feel under any pressure to rush. Take as much time as you need.

Please note

- No identifying information about participants will be published in any form
- You are free to withdraw from the study at any time and without giving reasons for withdrawing.

A note about virtual reality equipment

When using projection systems, people sometimes experience a degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please say so and we will stop.

With any type of video equipment there is a possibility that an epileptic episode may be generated. This, for example, has been reported for computer and video games. For this reason we regret that we are unable to accept volunteers who are known to have suffered from epilepsy.

Please turn over for the experiment consent form.....

Consent Form

Project: PhD Project – Socially Situated Simian Simulation

To be completed by volunteers: We would like you to read the following questions carefully.

Have you read the information sheet about this study?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all your questions?	YES/NO
Have you received enough information about this study?	YES/NO
Which investigator have you spoken to about this study? _____	
Do you understand that you are free to withdraw from this study?	
• At any time	YES/NO
• Without giving a reason for withdrawing	YES/NO
Do you agree to take part in this study?	YES/NO
Do you agree to be video taped?	YES/NO
Do you agree to be audio taped?	YES/NO
Do you agree to your heart rate being recorded?	YES/NO
Do you agree to your cognitive load being recorded?	YES/NO
Do you agree to the data being stored and used in the researcher's ongoing research?	YES/NO
Do you agree to the data being stored indefinitely?	YES/NO
Do you agree to let us make transcripts of your interactions and present them in printed publications?	YES/NO
Do you agree to let us make transcripts of your interactions and present them in presentations?	YES/NO
Do you agree to let us use audio fragments involving you in presentations?	YES/NO
Do you agree to let us use video frames or photo stills of you in printed publications?	YES/NO
Do you agree to let us use the video or photo stills of you in presentations?	YES/NO
Do you agree to the audio, video and other data being shared with academic collaborators on the project at University of Salford?	YES/NO
Do you agree to completing a personality questionnaire?	YES/NO

I certify that I do not have epilepsy.

Signed.....Date.....

Name in block letters.....

Investigator.....

In case you have any enquiries regarding this study in the future, please contact:

Simon Campion
The Centre for Virtual Environments and Future Media,
THINKLab,
7th Floor Maxwell Building,
University of Salford,
Salford, Greater Manchester, M5 4WT
+44 161 295 2931
Email: s.p.campion@salford.ac.uk

Appendix 7: Ethical Approval for Experiment 2

Academic Audit and Governance Committee

College of Science and Technology Research Ethics Panel
(CST)

University of
Salford
MANCHESTER

To Simon Campion (and Prof David Roberts)
cc: Prof Sunil Vadera, Head of School of CSE
From Nathalie Audren Howarth, College Research Support Officer
Date 26 November 2013

MEMORANDUM

Subject: Approval of your Project by CST
Project Title: Socially Situated Simian Simulation
REP Reference: CST 13/117

Following your responses to the Panel's queries, based on the information you provided, I can confirm that they have no objections on ethical grounds to your project.

If there are any changes to the project and/or its methodology, please inform the Panel as soon as possible.

Regards,

Nathalie Audren Howarth
College Research Support Officer

For enquiries please contact:
College of Science and Technology
College Research Support Officer
The University of Salford
Maxwell building, (7th floor, room 721)
Telephone: 0161 295 5278
Email: n.audren@salford.ac.uk