# Perception of Emotional Body Language Displayed by Animated Characters

PhD Thesis

School of Creative Technologies

University of Portsmouth

Aryel Beck

September 2011

The thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth.

# Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

# Abstract

Virtual Environments have demonstrated effectiveness for social task training such as medical training (Anolli, Vescovo, Agliati, Mantovani, & Zurloni, 2006). These types of Virtual Environments have used emotional animated characters. Even though emotions have a strong influence on human-human interactions (Gratch, Mao, & Marsella, 2006), typical system evaluation does not assess whether human and animated emotional displays are perceived similarly by observers. Moreover, the Uncanny Valley, which is a drop in believability as characters become more realistic, threatens the assumption that emotions displayed by an animated character and a human would be interpreted similarly. Thus, it is not known how appropriate the perception to a realistic emotional animated character is. This issue is especially important for social task training which require animated characters to be perceived as social and emotional partners so that trainees would be confronted with situations comparable to real life ones.

Using an approach similar to the one proposed by Nass & Moon (2000) in their work on the Media Equation, this thesis investigates how emotional body language displayed by animated characters is interpreted. A psychological experiment was conducted to investigate if emotional body language would be an appropriate way for animated characters to display emotion. This was done by comparing the interpretation of emotional body language displayed by animated characters with that by real actors. The results showed that animated body language can be accurately interpreted. However, the videos of the actor were found to be more emotional, more believable and more natural than the animated characters, whilst displaying the same emotional body language. Moreover, there was a significant difference in the number of correctly interpreted negative emotions displayed. Although, there was not a difference for positive emotions. This could be due to the physical appearance of the animated character or to the loss of micro-gestures inherent to Motion Capture technology.

Thus, a second comparative study was conducted to investigate the potential causes for this drop in believability and recognition. It investigated the effect of changing the level of physical realism of the animation as well as deteriorating the quality of the emotional body language itself. Whilst no effect was found regarding the deterioration of the emotional body language, the results show that the videos of the Actor were found to be more emotional, more believable and more natural than the two animated characters. These findings have strong implications for the use of Virtual Environments for social task training.

# Acknowledgments

I am grateful to Brett Stevens, my supervisor, for his infallible support, encouragements and guidance throughout the whole Ph.D. I am also indebted to Kim A. Bard for her guidance and support.

I am also thankful to Christina Howell-Richardson who provided advice and guidance helping me shaping ideas from the beginning.

I am also thankful to Arnaud Ducamp for reading the manuscript providing both corrections and critique.

I would like to take the opportunity to thanks Lola Cañamero for giving me the chance to continue the strand of research developed throughout this thesis.

I am also thankful to my colleagues and friends at the School of Creative Technologies: Steve Hand, Wendy and Vaughan Powell, Corinne Collier, Neil Dansey, Mitu Kandhaker, Si Qiao, Ruiying Wang, Roger Eglin, Terry Carnell and all the other members of staff.

I am deeply thankful to Thu Ha Nam for her love and constant encouragements throughout these years.

Last but not least, I am heartfelt thankful to my grandparents Annie and Pola, to my aunt Hélène, to my parents Cornelya and Andre, and to my brothers Yoav, Romain and Alexy-Assaf for their love and support. I am also thankful to Shino and Sarah, who recently joined the family.

# Related Publications

Beck, A., Stevens, B., Howell-Richardson, C., & Kalus, T. (2007). *Toward realistic emotional characters for training health professional.* Poster presented at Cybertherapy 2007, Washington DC.

Beck, A. (2007). *Realistic simulation of emotion by animated characters.* In proceedings of the Doctoral consortium, ACII 2007, Lisbon, Portugal.

Beck, A., Stevens, B., & Bard, K. (2008). Extending the Media Equation to Emotions: An Approach for Assessing Realistic Emotional Characters. *Annual Review of Cybertherapy and Telemedicine, 6*, 29-33.

Beck, A., Stevens, B., & Bard, K. (2009). *Comparing perception of affective body movements displayed by actors and animated characters.* In proceedings of the Symposium on Mental States, Emotions and their Embodiment, AISB09, Edinburgh, UK.

Beck, A., Stevens, B, Bard, K. & Cañamero L. (2012). Emotional Body Language Displayed by Artificial Agents. *ACM Transactions on Interactive Intelligent Systems, 2(1).*

# Contents


Declaration ....... iii
Abstract ....... iv
Acknowledgments ....... v
Related Publications ....... vi
Contents ....... vii
List of Figures ....... x
List of Tables ....... xi

**CHAPTER 1 - INTRODUCTION ....... 1**

1.1. Training and Virtual Environments ....... 1
1.2. Presence ....... 3
1.3. Social Task Training using Virtual Environments ....... 7
1.4. Body-Language as a Modality to Display Emotions Realistically ....... 11
1.5. Issue with the use of Virtual Environments for Social Task Training ....... 14
1.6. Thesis Statement and Overview ....... 17
*1.6.1. The Research Question ....... 17*
*1.6.2. Contribution to Knowledge ....... 19*
1.7. Structure of the Document ....... 21

**CHAPTER 2 - INVESTIGATING THE UNCANNY VALLEY ....... 23**

2.1. Introduction ....... 23
2.2. Definition and Context of the Uncanny Valley ....... 24
2.3. Empirical Evaluations of the Uncanny Valley ....... 27
*2.3.1. Research supporting the existence of the Uncanny Valley ....... 27*
*2.3.2. Research contradicting the existence of the Uncanny Valley ....... 28*
2.4. Potential causes of the Uncanny Valley ....... 29
2.5. Conclusion ....... 31

**CHAPTER 3 - BODY LANGUAGE AS A WAY FOR ANIMATED CHARACTERS TO DISPLAY EMOTIONS ....... 34**

3.1. Introduction ....... 34
3.2. The Media Equation ....... 34
3.3. Extending the Media Equation to Emotion ....... 36
3.4. The Experiment ....... 38
*3.4.1. Design ....... 38*
*3.4.2. Participants ....... 40*
*3.4.3. Apparatus ....... 40*
*3.4.4. Materials ....... 43*
*3.4.5. Procedure ....... 45*
3.5. Results ....... 45
*3.5.1. Data Confirmation ....... 45*
*3.5.2. Physical Realism ....... 47*
*3.5.3. Number of correct interpretations ....... 48*
*3.5.4. Emotional Strength ....... 54*
*3.5.5. Believability ....... 56*
*3.5.6. Naturalness ....... 57*
*3.5.7. Individual Differences ....... 57*
3.6. Discussions ....... 59
*3.6.1. Interpretations of the Displays ....... 59*
*3.6.2. Effect of Character Type on Strength, Believability and Naturalness ....... 60*
*3.6.3. Individual Differences are related to the perception of the displays ....... 61*
3.7. Conclusion ....... 62

**CHAPTER 4 - A COMPARATIVE STUDY OF EMOTIONAL BODY LANGUAGE DISPLAYED BY HUMANS AND ANIMATED CHARACTERS: STUDY DESIGN ........ 64**

4.1. INTRODUCTION ........ 64
*4.1.1. Secondary Animation ........ 64*
*4.1.2. Visual Realism ........ 64*
4.2. METHOD ........ 65
*4.2.1. Design ........ 65*
*4.2.2. Independent Variables ........ 66*
*4.2.3. Dependent Variables ........ 67*
*4.2.4. Research Questions ........ 68*
*4.2.5. Participants ........ 69*
*4.2.6. Apparatus ........ 69*
*4.2.7. Materials ........ 73*
*4.2.8. Procedure ........ 74*

**CHAPTER 5 - A COMPARATIVE STUDY OF EMOTIONAL BODY LANGUAGE DISPLAYED BY HUMANS AND ANIMATED CHARACTERS: STUDY RESULTS ........ 75**

5.1. DATA CONFIRMATION ........ 75
5.2. OVERVIEW OF THE EXPERIMENTAL EFFECTS ........ 76
5.3. EFFECTS OF FRAME RATE ........ 77
*5.3.1. Effect of Frame Rate on Correct Identification ........ 77*
*5.3.2. Effect of Frame Rate on Strength ........ 78*
*5.3.3. Effect of Frame Rate on Believability ........ 79*
*5.3.4. Effect of Frame Rate on Naturalness ........ 80*
*5.3.5. Summary of the effects of Frame Rate ........ 80*
5.4. REINVESTIGATION OF THE EFFECTS OF CHARACTER TYPE ON CORRECT IDENTIFICATION, STRENGTH, BELIEVABILITY AND NATURALNESS ........ 81
5.5. EFFECT OF CHARACTER TYPE ........ 83
*5.5.1. Effect of Character Type on Correct Identification ........ 83*
*5.5.2. Effect of Character Type on Strength ........ 86*
*5.5.3. Effect of Character Type on Believability ........ 87*
*5.5.4. Effect of Character Type on Naturalness ........ 88*
*5.5.5. Summary of the effects of Character Type ........ 89*
5.6. REINVESTIGATION OF THE EFFECTS OF ACTION STYLE ........ 90
5.7. INTERACTIONS BETWEEN ACTION STYLE AND CHARACTER TYPE. ........ 92
5.8. EFFECT OF ACTION STYLE ........ 95
*5.8.1. Effect of Action Style on Correct Identification ........ 95*
*5.8.2. Effect of Action Style on Strength ........ 96*
*5.8.3. Effect of Action Style on Believability ........ 97*
*5.8.4. Effect of Action Style on Naturalness ........ 98*
*5.8.5. Summary of the effects of Action Style ........ 98*
5.9. ANALYSIS OF INDIVIDUAL DIFFERENCES ........ 99

**CHAPTER 6 - A COMPARATIVE STUDY OF EMOTIONAL BODY LANGUAGE DISPLAYED BY HUMANS AND ANIMATED CHARACTERS: DISCUSSION OF THE RESULTS ........ 101**

6.1. INTRODUCTION ........ 101
6.2. CHARACTER TYPE DOES NOT AFFECT THE CORRECT IDENTIFICATION OF THE EMOTIONAL BODY LANGUAGE DISPLAYED. 101
6.3. CHARACTER TYPE AFFECTS THE STRENGTH OF THE EMOTIONAL BODY LANGUAGE DISPLAYED. ........ 103
6.4. CHARACTER TYPE AFFECTS THE BELIEVABILITY AND NATURALNESS OF THE EMOTIONAL BODY LANGUAGE DISPLAYED. 105
6.5. ACTION STYLE AFFECTS THE CORRECT IDENTIFICATION, STRENGTH, BELIEVABILITY AND NATURALNESS OF THE EMOTIONAL BODY LANGUAGE DISPLAYED. ........ 105
6.6. MOVEMENTS' SMOOTHNESS DOES NOT AFFECT THE CORRECT IDENTIFICATION, BELIEVABILITY AND NATURALNESS OF THE EMOTIONAL BODY LANGUAGE DISPLAYED. ........ 106
6.7. INDIVIDUAL DIFFERENCES ARE RELATED TO CORRECT IDENTIFICATION, STRENGTH, BELIEVABILITY AND NATURALNESS. 107

6.8. CONCLUSION ........ 108

**CHAPTER 7 – CONCLUSIONS ........ 109**

7.1. RESEARCH FINDINGS ........ 109
7.2. IMPLICATIONS FOR SOCIAL TASK TRAINING ........ 110
7.3. IMPLICATIONS FOR THE UNCANNY VALLEY ........ 111
7.4. IMPLICATIONS FOR PSYCHOLOGY ........ 113
7.5. CONFIRMATION OF RESULTS WITH ROBOTICS ........ 114
7.6. RECOMMENDATION FOR FUTURE WORK ........ 116
*7.6.1. Effect of the loss of believability on Social Presence ........ 116*
*7.6.2. Other Aspects of Realism ........ 117*
7.7. FINAL COMMENT ........ 118

**BIBLIOGRAPHY ........ 119**

**APPENDIX A ........ 126**

A.1. ISSUES WITH THE PRINCIPLES OF ANIMATION ........ 126
*A.1.1. Traditional animation principles that can be used to increase realism ........ 127*
*A.1.2. Traditional animation principles that cannot be applied to social task training ........ 130*

# List of Figures


**Figure 1.1** Technical elements required for accurate transfer of training from virtual to real world .................... 2
**Figure 1.2** Technical and psychological elements required for accurate transfer of training from virtual to real world ........ 4
**Figure 1.3** Technical, psychological and emotional elements required for accurate transfer of training from virtual to real world. ........ 5
**Figure 1.4** – Mission rehearsal Exercise and, FearNot........ 9
**Figure 1.5** Elements specific to the animated characters within a Virtual Environment. ........ 10
**Figure 1.6** Elements specific to the animation of body language........ 12
**Figure 1.7 :** The Uncanny Valley (adapted from Mori 1970) ........ 14
**Figure 1.8** Elements required for accurate transfer of skills lead to the Uncanny Valley ........ 16

**Figure 2.1 :** The Uncanny Valley (adapted from Mori 1970) ........ 24
**Figure 2.2:** Screen shot from Polar Express and from Final Fantasy ........ *25*
**Figure 2.3** The conditions necessary for accurate transfer of skills are potentially similar to those included in the Uncanny Valley. ........ 26
**Figure 2.4:** Example of faces used by MacDorman and Ishiguro (2006) ........ 27
**Figure 2.5:** Example of faces used by Hanson (2006)........ 28
**Figure 2.6** : Kaspar, an Humanoid Robot for Human-Robot Interaction Research ........ 29

**Figure 3.1** Emotional Perception in the context of Virtual Environments........ 36
**Figure 3.2** Steps to create the experiment's material ........ 40

**Figure 3.3.** Screen shots of video condition and animated condition ........ 40, 42
**Figure 3.4.** The Geneva Emotional Wheel (Scherer, 2005)........ 44
**Figure 3.5.** Physical Realism of the Actor and of the Animated Character........ 47
**Figure 3.6.** Means for *Correct Interpretation* ........ *49*
**Figure 3.7.** Means for *Correct Interpretations* for each *Character Type, Action Style* and *Context*. ........ 50
**Figure 3.8.** Emotional *Strength* for each *Character Type*, ........ 54
**Figure 3.9.** Emotional *Strength* for each *Character Type*, *Action Style* and *Context* ........ *55*
**Figure 3.10.** *Believability* of the Actor and of the Animated Character........ 56
**Figure 3.11.** *Naturalness* of the Actor and of the Animated Character........ 57

**Figure 4.1.** Screen shots of the Actor, of the Realistic Character and of the Simplified Character. ........ 70
**Figure 4.2.** Screen shot of the application........ 73
**Figure 4.2.** Photo of the experimental setting........ 74

**Figure 5.1.** The overall *Strength* by *Character Type* ........ 81
**Figure 5.2.** The overall *Believability* by *Character Type*........ 82
**Figure 5.3.** The overall *Naturalness* by *Character Type* ........ 82
**Figure 5.4.** The psychological space of interpretation of emotions display ........ 85
**Figure 5.5.** The Overall *Correct Identification* for the two styles........ 90
**Figure 5.6.** The Overall *Strength* for the two styles........ 90
**Figure 5.7.** The Overall *Believability* for the two styles ........ 91
**Figure 5.8.** The Overall *Naturalness* for the two styles ........ 91
**Figure 5.9.** The overall *Strength* for the two styles for each character. ........ 92
**Figure 5.10.** The overall *Believability* for the two styles for each character. ........ 93
**Figure 5.11.** The overall *Naturalness* for the two styles for each character. ........ 94

**Figure 7.1** Elements specific to the animated characters within a Virtual Environment ........ 110
**Figure 7.2** Three possible positions of the characters tested on the Uncanny Valley........114
**Figure 7.3** The six key poses ........ 115
**Figure 7.4** Elements required for accurate transfer of skills lead to the Uncanny Valley. ........ 116
**Figure A.1:** Screen shot from Snow White (Disney 1937) ........ *126*

# List of Tables


**Table 3.1**: Summary of the Media Equation results. ............ 35
**Table 3.2**. Experiment conditions. ............ 39
**Table 3.3.** Means and Standard deviations of the results. ............ 46
**Table 3.4.** Proportion of good answers per emotions and conditions ............ 48
**Table 3.5.** Confusion Matrix of the Acted Ordinary displays ............ 51
**Table 3.6.** Confusion Matrix of the Animated Ordinary displays ............ 52
**Table 3.7.** Confusion Matrix of the Acted Stylised displays ............ 52
**Table 3.8.** Confusion Matrix of the Animated Stylised displays ............ 53

**Table 4.1.** Experiment conditions ............ 66

**Table 5.1.** Percentage of participants that identified correctly the emotion at least once ............ 75
**Table 5.2.** Results of the ANOVAs carried out on *Correct Identification* between *Frame Rate* for each Emotion Displayed ............ 77
**Table 5.3.** Results of the ANOVAs carried out on *Strength* between *Frame Rate* for each Emotion Displayed . 78
**Table 5.4.** Results of the ANOVAs carried out on *Believability* between *Frame Rate* for each Emotion Displayed ............ *79*
**Table 5.5.** Results of the ANOVAs carried out on *Naturalness* between *Frame Rate* for each Emotion Displayed ............ *80*
**Table 5.6.** Results of the Repeated Measures ANOVAs carried out on *Correct Identification* between *Character Type* for each Emotion Displayed. ............ 83
**Table 5.7.** Results of the Repeated Measures ANOVAs carried out on *Strength* between *Character Type* for each Emotion Displayed ............ *86*
**Table 5.8.** Results of the repeated measures ANOVA and post-hoc tests carried on *Believability* between *Character Type* ............ *87*
**Table 5.9.** Results of the repeated measures ANOVA and post-hoc tests carried on *Naturalness* between *Character Type* ............ 88
**Table 5.10.** Results of the repeated measures ANOVA carried out on *Correct Identification* between *Action Style* for each emotion displayed. ............ 95
**Table 5.11.** Results of the repeated measures ANOVA carried out on *Strength* between *Action Style* for each emotion displayed ............ *96*
**Table 5.12.** Results of the repeated measures ANOVA carried out on *Believability* between *Action Style* ........ 97
**Table 5.13.** Results of the repeated measures ANOVA carried out on *Naturalness* between *Action Style* ........ *98*

# Chapter 1 - Introduction

## 1.1. Training and Virtual Environments

Training, the action of teaching or learning particular skills, is typically done by instruction and practice. However, experiential learning theory places experiences at the centre of the learning process (Jarmon, Traphagan, Mayrath, & Trivedi, 2009). According to Kolb's experiential learning theory, learning occurs through "concrete experiences", which provides the basis for "reflective observations". In turn these observations become "abstract concepts", which can be tested through "active experimentation" (Kolb, 1984). In this model, instruction works at the "reflective observation" and the "abstract conceptualization" levels. It leads trainees to capitalise and reflect upon their experiences, helping them to turn these into concepts. Practice works at the "concrete experience" and the "active experimentation" levels. It helps trainees experience the field and test the concepts they learned. Moreover, depending on the tasks, this can be done either in the real situation with a tutor or in a simulation, e.g. role-playing, and can include the training of social and contextual skills.

In addition to traditional techniques, such as role playing, Virtual Environments are also becoming increasingly used for training. They provide instructors with a tool that can also work at the concrete experience and active experimentation levels by producing simulated experiences. Moreover, as computer graphics improve, Virtual Environments can increasingly portray more complex and realistic situations. Indeed, the technology has passed the stage in which it is possible to create Virtual Worlds that simulate environments that act under the control of their users (Bartle, 2003). These are most commonly used in the computer games industry, which creates Virtual Environments where players embody different characters in simulated settings. However, along with entertainment examples, they have also been successfully used for 'serious games', which are used for training.

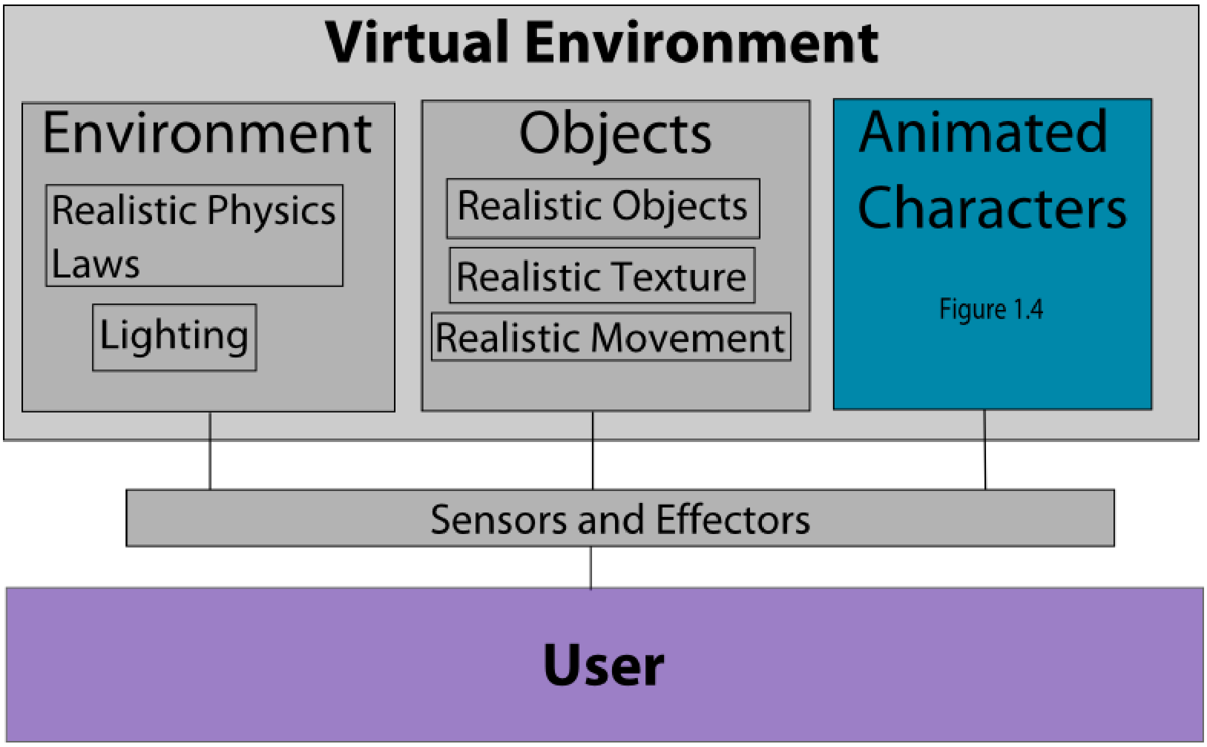


**Figure 1.1** Technical elements required for accurate transfer of training from virtual to real world. The coloured boxes will be developed throughout the thesis

When designed for training, Virtual Environments simulate real life situations in which trainees can learn to perform complex tasks in a 'risk free' space. However, to do so, Virtual Environments need to be comparable to real life situations. Hence, they typically respect physics laws (such as gravity etc..) and use realistic lighting. Virtual Environments react to user's actions through sensors and effectors (Figure 1.1). Moreover, objects within a virtual environment, eventually including animated characters, must look as if they were real as well. This is done through the object's shape, colours (i.e. texture) and movements (Figure 1.1). For example, Vora et al. (2002) developed a Virtual Environment for aircraft visual inspection training. They developed a library of defects for different aircrafts that can be found during a real inspection. These include corrosion and cracked & broken conducts. In this Virtual Environment, trainees inspect computer-generated aircrafts following the same procedures as with real aircrafts. By practising inspection in the Virtual Environment, trainees learn the procedures that need to be followed (Vora et al., 2002). Their findings show that their system could improve the quality of the training provided for such inspections (Vora, et al., 2002) and that such systems are preferred over more 'classic' computer based training systems (Vora, et al., 2002). Moreover, any missed defects do not endanger lives of real crew or passengers.

Virtual Environments have also been created for the training of medical students to perform specific operations. Recently, a Virtual Environment developed for surgical training enables trainees to practise a complex operation on a cataract (phacoemulsification) (Choi, Soo, & Chung, 2009). In this environment, trainees can acquire the experience and the dexterity required to perform this complex eye operation at no risk to a patient.

Using Virtual Environments, trainees can learn to perform operations in a low risk setting. However, Virtual Environments for training have focused on mechanical tasks, (e.g. Vora, et al., 2002 and Choi, et al.,2009), as these are relatively simple to create. To date, most of the Virtual Environments that are used for training are comparable to these, i.e. they are task specific and trainees learn to perform 'mechanical' tasks or to follow pre-defined procedures. Moreover, they have been found to be effective for this type of training (Darken & Banker, 1998; Grantcharov, Kristiansen, Bendix, Bardram, & Funch-Jensen, 2003; Rose, Attree, Brooks, Parslow, & Penn, 2000).

However, Virtual Environments are now also starting to be developed for social task training, see for example (Zhang et al., 2009). Highly emotive and high human cost tasks, such as medical diagnosis, could benefit from using Virtual Environments. These could enrich training with non mechanical experiences such as conversational skills that are essential to learn prior to being confronted with real situations. However, this introduces an additional layer of complexity over simple machine fault diagnosis, as it is necessary to simulate the way humans interact with each other. Moreover, the effectiveness of this has not been clearly established. It is in this context that this thesis is based.

## 1.2. Presence

The learning potential of Virtual Environments relies on the possibility for learners to take part in a number of experiences, in a safe and protected setting. In order to be effective for training, Virtual Environments need to provide accurate transfer of training. Transfer of training is the ability to apply, in real life, what has been learned in the Virtual Environment (Hamblin, 2005). However, for this to occur the experience should seem real and engaging to participants, they should 'feel' present in the situation (Mantovani & Castelnuovo, 2003). Thus, one of the key requirements of Virtual Environments for training, especially social task training, is that they should induce a feeling of Presence.

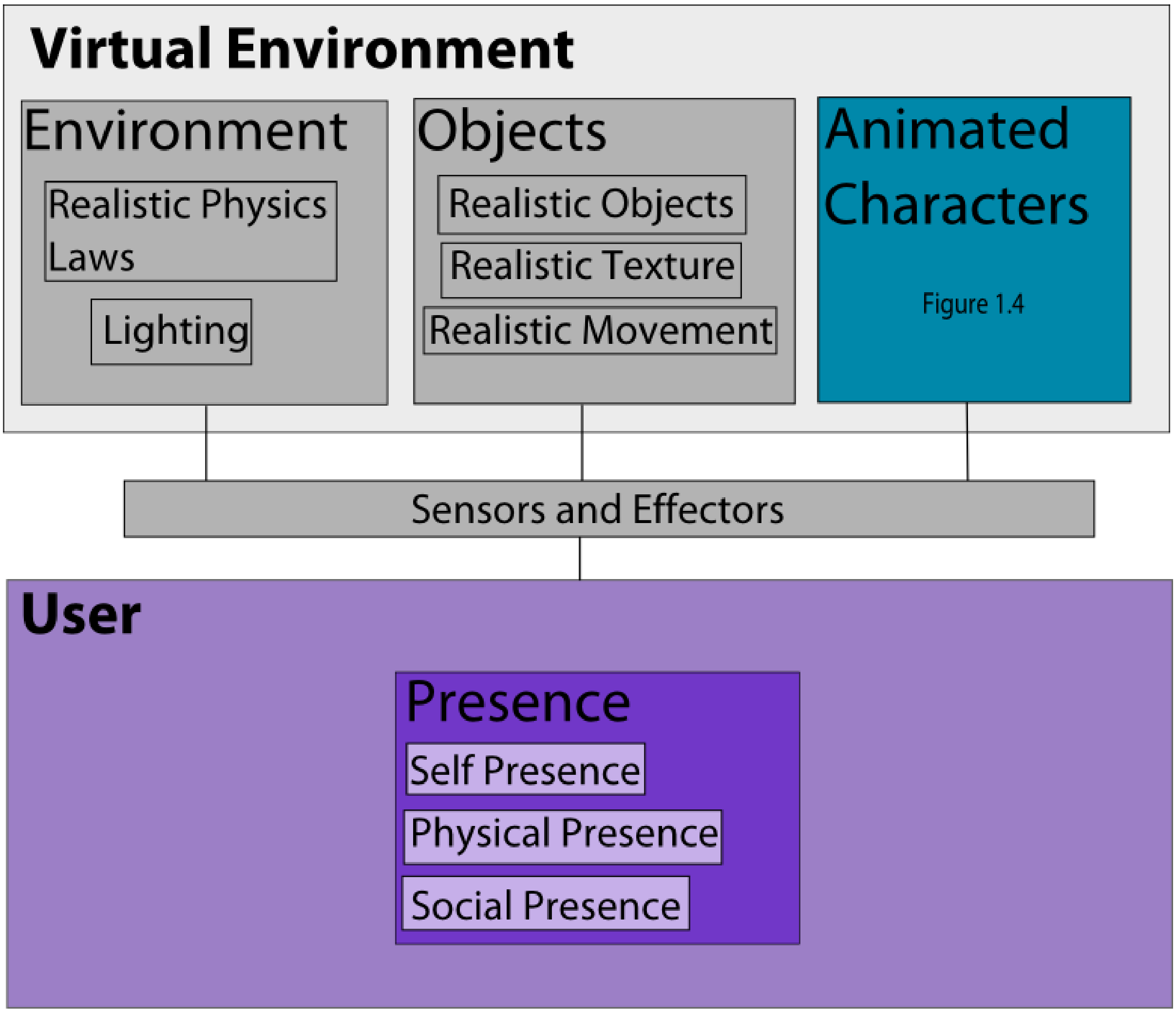


**Figure 1.2** Technical and psychological elements required for accurate transfer of training from virtual to real world.

*"The concept of 'presence' refers to the phenomenon of behaving and feeling as if we are in the virtual world created by computer displays"* (Sanchez & Slater, 2005). In virtual reality presence is generally studied along three main components: 'Physical' presence, 'Social' presence and 'Self' presence (Heeter, 1992; Lee, 2004). 'Physical' presence is defined as a psychological state in which virtual objects are experienced as actual physical objects in either sensory or non-sensory ways (Lee, 2004). It is the perception of objects as 'real' in the environment, which leads to a greater sense of 'being there' (Figure 1.2).

'Social' presence is defined as a psychological state in which virtual characters are experienced as actual social actors in either sensory or non-sensory ways (Lee, 2004). In other words, if others (computer generated agents or other human beings present in the virtual environment) engage socially with the user, and take into account the user's presence when acting in the Virtual Environment, it reinforces the physical sensation of being there (Figure 1.2).

Finally, ‘Self’ presence is defined as a psychological state in which virtual self/selves is/are experienced as the actual self in either sensory or non-sensory ways (Lee, 2004). In other words it is the perception that one has one’s own body in the environment, i.e. the virtual body replaces the ‘real’ body as the primary ‘effector’, the sense of self moves from ‘here’ to ‘there’.

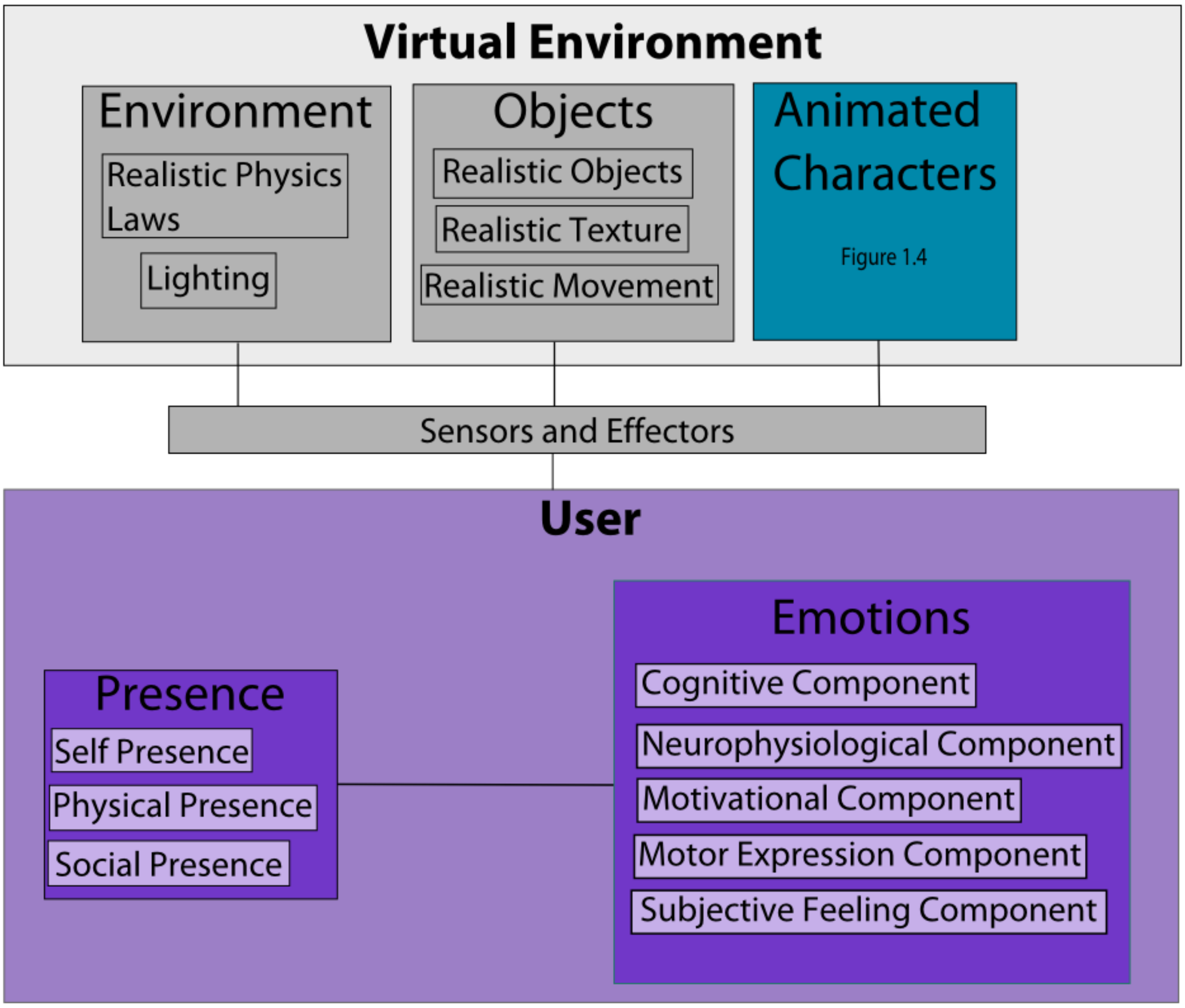


**Figure 1.3** Technical, psychological and emotional elements required for accurate transfer of training from virtual to real world.

Nevertheless, such definitions are limited because they reduce presence to cognition, ignoring the emotional aspect of presence (Banos et al., 2004). This is reinforced by Huang & Alessi (1999) who argue that “*any theory of presence must take emotional factors into account*”. Moreover, a sign of presence in a Virtual Environment is to exhibit behaviour similar to what the users' behaviour would have been in a comparable real life situation (Slater, 2003). One may therefore suggest that when someone 'feels' present in an environment they react emotionally to it as they would in the real world (Figure 1.3). Moreover, it could also be argued that presence is necessary for users to react emotionally as well as to perceive emotions in others (Figure 1.3).

According to Scherer (1987, 2005), an emotion can be defined as *"an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism"*. In this theoretical model, the Cognitive Component (Figure 1.3) is responsible of the evaluation of the stimulation. The Neurophysiological Component (Figure 1.3) is responsible for system regulation. The Motivational Component (Figure 1.3) affects preparation and direction of coming actions. The Motor Expression Component (Figure 1.3) communicates intent and reactions. Finally, the Subjective Feeling Component (Figure 1.3) is responsible for monitoring, attention focusing, and for reflection (Scherer, 1987, 2005).

There is empirical evidence supporting the idea of presence as a necessary condition for eliciting emotions related to the virtual environment. For example, using a virtual park (Riva et al., 2007) sought to establish whether "*the virtual environment is able to produce the feeling of presence in users, [and if] this environment will also be able to elicit emotions*." They found strong support for their theory, however, the experiment included a possible confound. The different music used in conjunction with the virtual park environment could have reinforced the intended emotions. The precise contribution of music to the induction of the mood was not established. Yet music has already been used in psychology research as a mood inducer (Kalat & Shiota, 2007), it may be the factor responsible for a mood change. The extent to which the feeling of presence in this virtual park is influenced by the music should have been investigated as well.

The relation between presence and emotion is of particular interest for social task training as these tasks have a strong emphasis on 'social' presence, which requires trainees to behave and interact with virtual characters as if they were real people. In order to trigger real life behaviours in trainees the characters should behave naturally; displaying emotions as humans do. The interactions should be made as realistic as possible, thus, ensuring that the skills developed in the Virtual Environment can be applied in real life, i.e. accurate transfer of training. If the characters do not induce social presence, perhaps because of unnatural interactions, it will not trigger real reactions in trainees. This would result in either not developing the required skills, or them not being transferrable to the real world, i.e., failures in training.

There is however a growing body of evidence that shows that users tend to engage with virtual characters as if they were real (Reeves & Nass, 1996; Slater et al., 2006; Triola et al., 2006). For example, Slater et al.(2006) tested if a virtual character would be treated as a real person in an immersive Virtual Environment. They used a simulated version of Milgram's obedience experiment. In their experiment, participants were instructed to deliver 'electric' shock to a virtual learner when it made a mistake. They found that participants reacted to the situation as if it were real, even though they knew the character was computer generated (Slater, et al., 2006).

## 1.3. Social Task Training using Virtual Environments

When compared with 'mechanical' task training, the Virtual Environments necessary for social task training are more complex to develop. In order for accurate transfer of training to occur, the virtual environment should reproduce situations comparable to the ones that occur in real life. This can occur if the Virtual Environment supports the actions from users, for instance grasping an object or looking around (Figure 1.1). Moreover, the objects and entities within the Virtual Environment need to look and behave like real ones (Figure 1.3). For instance, if the trainees talk to an animated character, it should respond appropriately. Thus, it is necessary to model virtual characters with which trainees can interact, bond and feel concern for, all as if they were 'real'.

Consider, for example, training medical students to interview and communicate with patients, which is an important but difficult task (O'Sullivan, Chao, Russell, Levine, & Fabiny, 2008). Medical interview is usually defined as the primary task to establish the medical background of a patient. It is the basis for establishing a therapeutic relationship with the patient, ascertaining a clinical diagnosis, and implementing an appropriate treatment plan or health maintenance (Cassata, 1983).

The use of real patients for such training raises ethical and practical issues for trainees as well as for patients. Currently, medical schools often rely upon standardised patients (actors trained to portray patients with health problems) to train medical students. The use of standardised patients makes it possible to produce situations comparable to the ones that occur in real life for the purpose of training. It has been found to be very positive for students' learning (Konkle-Parker, Cramer, & Hamill, 2002). However, such training remains limited by the availability of actors, as well as by the time and money necessary to train them. Moreover, when used for assessment, standardised patients may induce a bias as there is no guarantee that the actors will perform equally well among different students.

Medical schools also use pre-recorded videos, which have also been efficiently used for medical interview training (Lane & Gottlieb, 2004). The advantage of video training compared with standardised patients is that it does not rely on actors being present with the trainees. Once the video material has been developed it can be used by instructors and students in portable and relatively flexible ways. However, such training is always limited to the set of situations that have been previously recorded, i.e. video training cannot adapt to situations different from the ones previously recorded and therefore cannot adapt easily to input from the user. Moreover, as it is not interactive video training, it may not be comparable to the use of actors in terms of creating an emotional response in the trainee. This makes it difficult for the trainee to learn how to cope and regulate their own emotional state using this medium.

In comparison, Virtual Environments could combine some of the benefits of these two methods. Virtual Environments do not rely on actors being present with trainees. They are consistent in terms of performances and could therefore be used for assessment. Moreover, they are flexible and can adapt to the trainees' reactions (given an appropriate Artificial Intelligence, which are also being developed (Swartout et al., 2006)). The technology for creating the animated characters that populate Virtual Environments has already been developed to such a point that they are no longer 'background' characters but rather considered as entities with whom trainees can interact (Parsons et al., 2008). Thus, medical interview is an example of training that could benefit from the use of Virtual Environments (Parsons, et al., 2008).

**Figure 1.4** – Mission Rehearsal Exercise (Top), FearNot (Bottom)

There are existing examples of social task training, including "the mission rehearsal exercise"; a Virtual Environment for training military personnel going to serve in Iraq (Swartout, et al., 2006). In this simulation, the trainee takes on the role of a Lieutenant in a unit that has just experienced an accident involving civilians during a mission (Figure 1.4 Top). The task is to decide upon a course of action and then give appropriate instructions to the Sergeant. Trainees are immersed in a stressful situation surrounded by realistic characters as they need to develop the skills necessary to make 'good' decisions, whilst interacting and negotiating.

A very different looking example is "FearNot", which is a Virtual Environment for educating children on the issues of bullying (Aylett, Paiva, Woods, Hall, & Zoll, 2005). "FearNot" uses cartoonish characters with which children bond and sympathise while watching bullying situations from different perspectives (Figure 1.4 Bottom). In this Virtual Environment, children have to make decisions on the course of action a bullied character is taking. The virtual environment adapts to the decisions made by the children, thus demonstrating the appropriateness of a course of action they may take when witnessing or being confronted with such situations in real life.

Although they vary greatly in their level of physical realism (Figure 1.4), both "the mission rehearsal exercise" and "FearNot" use 'emotional' animated characters and have demonstrated good learning outcomes through the portrayal of realistic scenarios. Such Virtual Environments are portable and flexible and aim at simulating situations comparable to the ones that occur in real life. However, they require detailed animated characters.

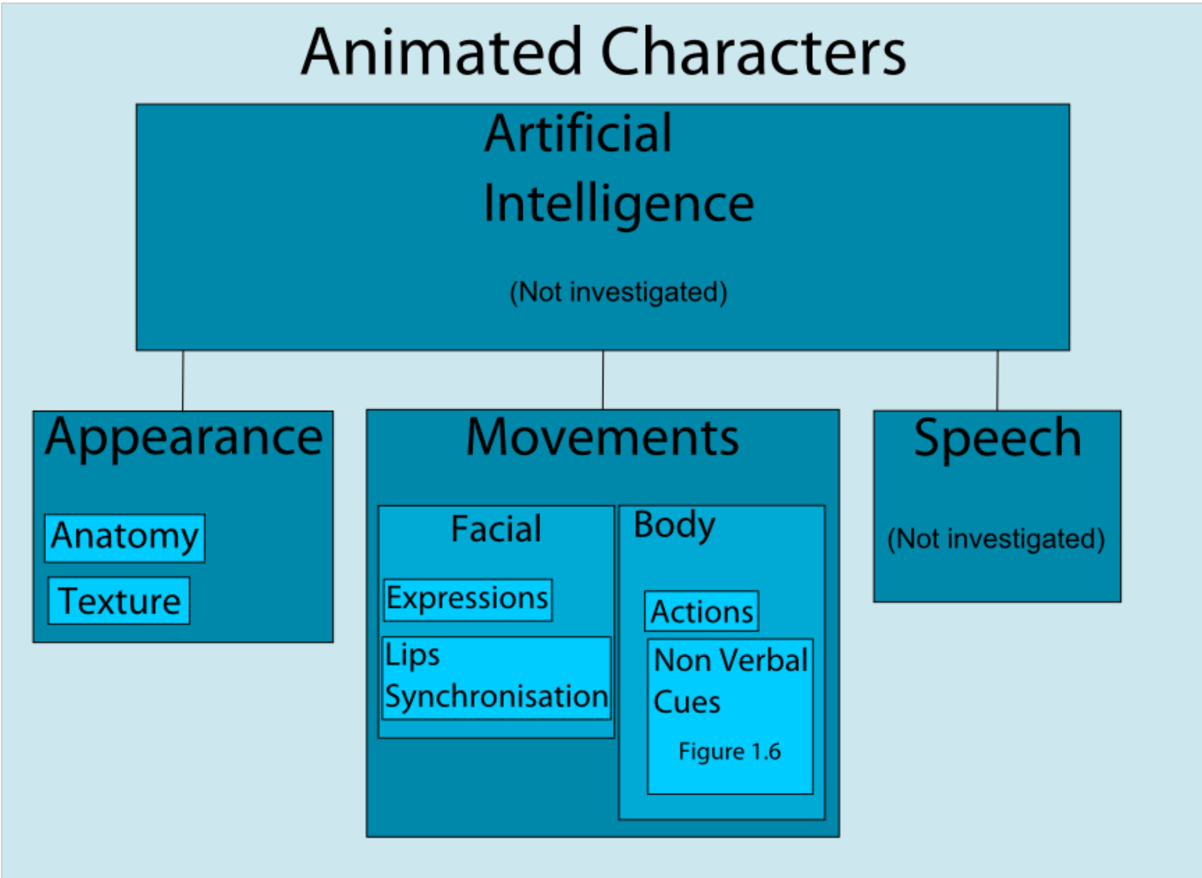


**Figure 1.5** Elements specific to the animated characters within a Virtual Environment.

Virtual Environments present a promising setting for social task training. This is one of the reasons why work toward the creation of animated characters, such as virtual patients, that model mental and physical states (Parsons, et al., 2008), is currently being investigated. This includes appropriate behavioural responses, for which the animated characters' need to be supported by artificial intelligence (Figure 1.5). For instance, Greta is a state of the art virtual human that can display non verbal expressions synchronised with speech as well as emotions using its face in combination with gestures (Pelachaud, 2009). Some existing Virtual Environment systems use computational models for emotions and actions (Aylett, et al., 2005; Swartout, et al., 2006). However, the artificial intelligence aspects are beyond the scope of this thesis (Figure 1.5).

In addition to object collision, texture, etc., work is also beginning on simulating a wider set of human interactions (Vinayagamoorthy et al., 2006) increasing the realism of the animated characters.

Social task training, such as medical interview, require realistic characters. Without considering Artificial Intelligence, their level of realism can be defined along two main aspects: Physical Realism and Behavioural Realism. Physical Realism is related to the character appearance (Figure 1.5), it includes anatomy and texture (Figure 1.5). On the other hand, Behavioural Realism, which is related to the ways a character moves (Figure 1.5), includes a full range of simulated behaviours and actions, including speech, lip synchronisation, as well as animated facial and body expressions (Figure 1.5).

These aspects have started to be researched from a functional perspective (Vinayagamoorthy, et al., 2006). However, an animated character may also require more subtle behaviours such as non-verbal cues (Figure 1.5), or body language (Martin, Niewiadomski, Devillers, Buisine, & Pelachaud, 2006). This would be essential, for example, in training medical personnel in assessing a patient's inner emotional state (Hall, Harrigan, & Rosenthal, 1995). It should be noted though, that the success of such realistic cues within a Virtual Environment will always depend on how they are perceived and interpreted by the user. However, it is not evident that emotional expressions displayed by a human and by an animated character are understood by the viewer in a similar way.

## 1.4. Body-Language as a Modality to Display Emotions Realistically

To most fully support interpretation of emotional state, traditional animation suggests that focusing on the display of emotion through the body would be an appropriate approach (Appendix A). However, when it comes to emotional expressions, psychology has mainly focused on understanding facial expressions (de Gelder, 2006). Nevertheless, an animated character displaying emotion realistically through the face and not through the rest of the body will probably still look unnatural to a viewer. This has been codified as a rule in classical animations as "*the [emotional] expression must be captured throughout the whole body as well as in the face*" (Thomas & Johnston, 1995) (Appendix A). Theatre also does so, by asking actors to become, in Artaud's words, "athletes of the emotions" and a large part of an actor's training concerns the non-verbal expression of emotions (see AMS (2007) for example). Emotions such as fear, anger, happiness, stress… are likely to be expressed through the whole body and should be readable.

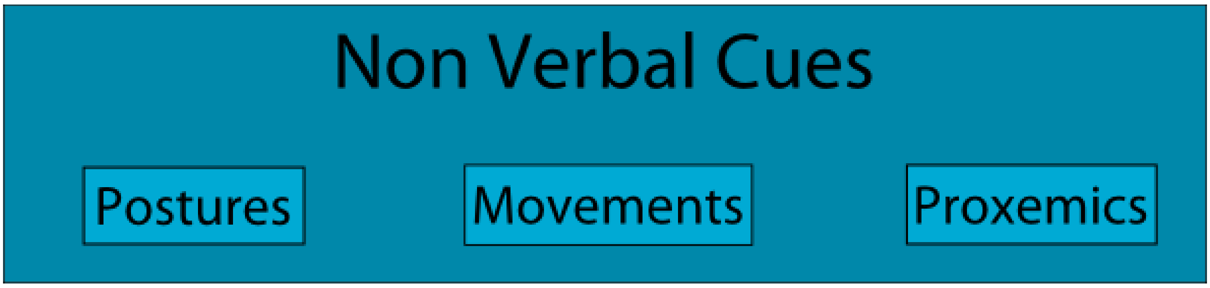


**Figure 1.6** Elements specific to the animation of body language.

Researchers have categorised the different types of body language, depending on how it occurs. The categorisation created from Vinayagamoorthy et al.(2006) and from Cassell (2000) can be separated into three different areas that should be considered when animating characters for social task training.

**Postures**: Postures are specific positions that the body takes during a timeframe. It has been established that postures are an effective medium to express emotion for humans. For instance, De Silva et al.(2004) investigated cross-cultural recognition of four emotions (anger, fear, happiness, sadness) through interpretations of body postures. They built a set using actors to perform emotional postures and showed that it was possible for participants to correctly identify the different emotions (De Silva & Bianchi-Berthouze, 2004). Thus, an animated human character displaying realistic emotion will also have to take up postures appropriate to the emotion (Figure 1.6).

**Movement**: It has been shown that many emotions are differentiated by characteristic body movements, and that these are effective clues for judging the emotional state of other people in the absence of facial and vocal clues (Atkinson, Dittrich, Gemmell, & Young, 2004). Thus, an animated character displaying emotions realistically should also do so during, and via, motion. Body movements include the movements themselves as well as the manner in which they are performed, i.e. movement speed, dynamics, curvature; Something captured by the traditional animation principles (Appendix A). Moreover, it should be noted that body movements occur in interaction with other elements such as speech, facial expressions, gaze, which would also be affected by the emotional state of the character. However, these elements are out of the scope of this thesis.

**Proxemics**: It is the distance between individuals during a social interaction. It is also indicative of emotional state. For example, angry people have a tendency to reduce the distance during social interaction, although this reduction would also be evident between intimate people. Proxemics cannot therefore be considered as an emotional expression in itself but is required to complete a representation of realistic emotional behaviour (Figure 1.6).

In order to display emotional body language comparable to a humans', these three categories should be considered for animated characters. However, the observers' ability to understand and react to the expressive cues displayed will have to be considered. There is no evidence that the cues displayed by an animated character and by a human would be interpreted in the same manner. This is an important issue, especially since displaying realistic body language conflicts with some of the principles of traditional animation (such as exaggeration or anticipation). It is therefore not clear that realistic animated characters would be efficient for training as they could result in a lack of believability, which would limit social presence (Figure 1.2). This is an important issue, as characters would have to be engaging in order for the simulation to induce social presence but would also have to produce a realistic experience comparable, to real life situations, in order to support accurate transfer of training (see Section 1.2 for definition).

Moreover, the use of Motion Capture technology, which records human motion that can be used to animate hyper-real characters, has shown that the 'real movements' do not always look correct, even though they are. This could be due to the fact that Motion Capture records joint positions and extrapolates the movements of the soft tissues from them, losing or procedurally generating secondary animations such as skin or muscle deformation (Section 3.6.2). However, the disregard of some of the traditional animation principles (such as exaggeration, squash and stretch...) could result in a lack of believability (i.e. the extent to which a character seems to be alive) in the characters (see Appendix A for further discussions).

# 1.5. Issue with the use of Virtual Environments for Social Task Training

Typically, animated character evaluation does not assess how the emotional displays are perceived by observers. Indeed, it is not evident that expressions displayed by a human and by an animated character are even interpreted in a similar way. However, if any differences were evident, it could exclude Virtual Environments as a tool for social task training. Encouraging results have however been found from much weaker stimulus. It was found that humans tend to interact with computers as they do with real people (Nass & Moon, 2000). These results were obtained using restricted technology and if it held true for realistic animated characters, it would suggest that social task training using Virtual Environment could be a promising alternative training to human-human interaction, a view supported by the Realism Maximisation theory (Groom et al., 2009).

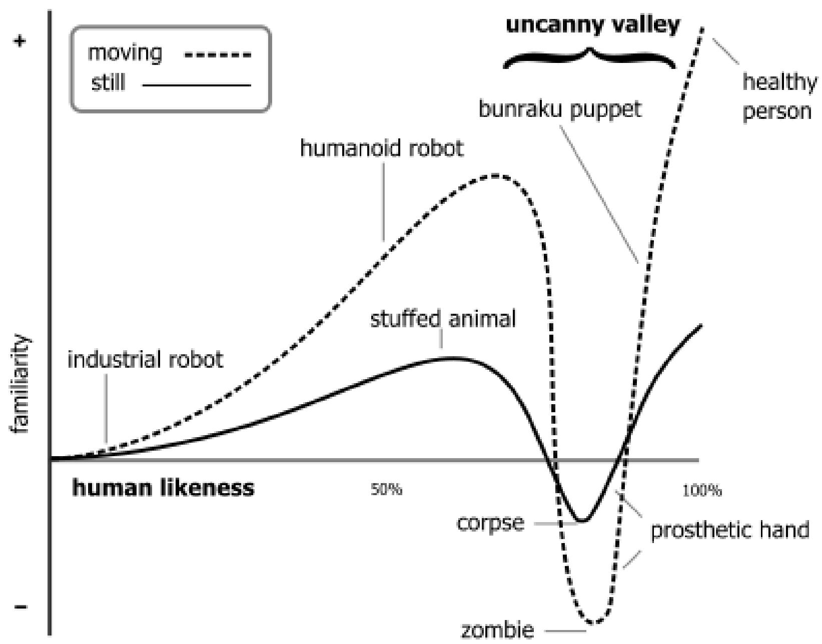


**Figure 1.7 :** The Uncanny Valley (adapted from Mori 1970) (Mori, 1970)

However, there is a problem with assuming that greater fidelity would improve the interaction. As animated characters become more realistic they have been confronted with the issue of the Uncanny Valley (Figure 1.7). The Uncanny Valley (Mori, 1970) models a drop in believability as characters acquire greater similarity with humans (Brenton, Gillies, Ballin, & Chatting, 2005). The concept was first introduced in robotics where it was reported that highly realistic humanoid robots tend to be found repulsive (Mori, 1970).

Animated characters from the film industry have been confronted with the same drop in believability as described by the Uncanny Valley (see Chapter 2 for further discussions). For example, the characters from “The Polar Express” (Zemeckis, 2004) or “Final Fantasy” (Sakaguchi & Sakakibara, 2001) have failed to be convincing (MacDorman, Green, Ho, & Koch, 2009). However, unrealistic characters, such as the ones used in traditional cel animation, are not confronted with this issue (Appendix A). They still generate empathy and to some extent they make the viewers ‘forget’ that they are not real. However, highly realistic characters that “fall in” the Uncanny Valley, fail to generate empathy, and constantly remind viewers that they are artificial.

Despite the anecdotal accounts, when applied to animation, the Uncanny Valley (Figure 1.7) is not supported by empirical results and its very existence is still subject to debate (Dautenhahn & Hurford, 2006). Indeed, the lack of credibility could result from poor technical design or disappointing storylines, or, as suggested by Brenton et al.(2005), from poor facial animation.

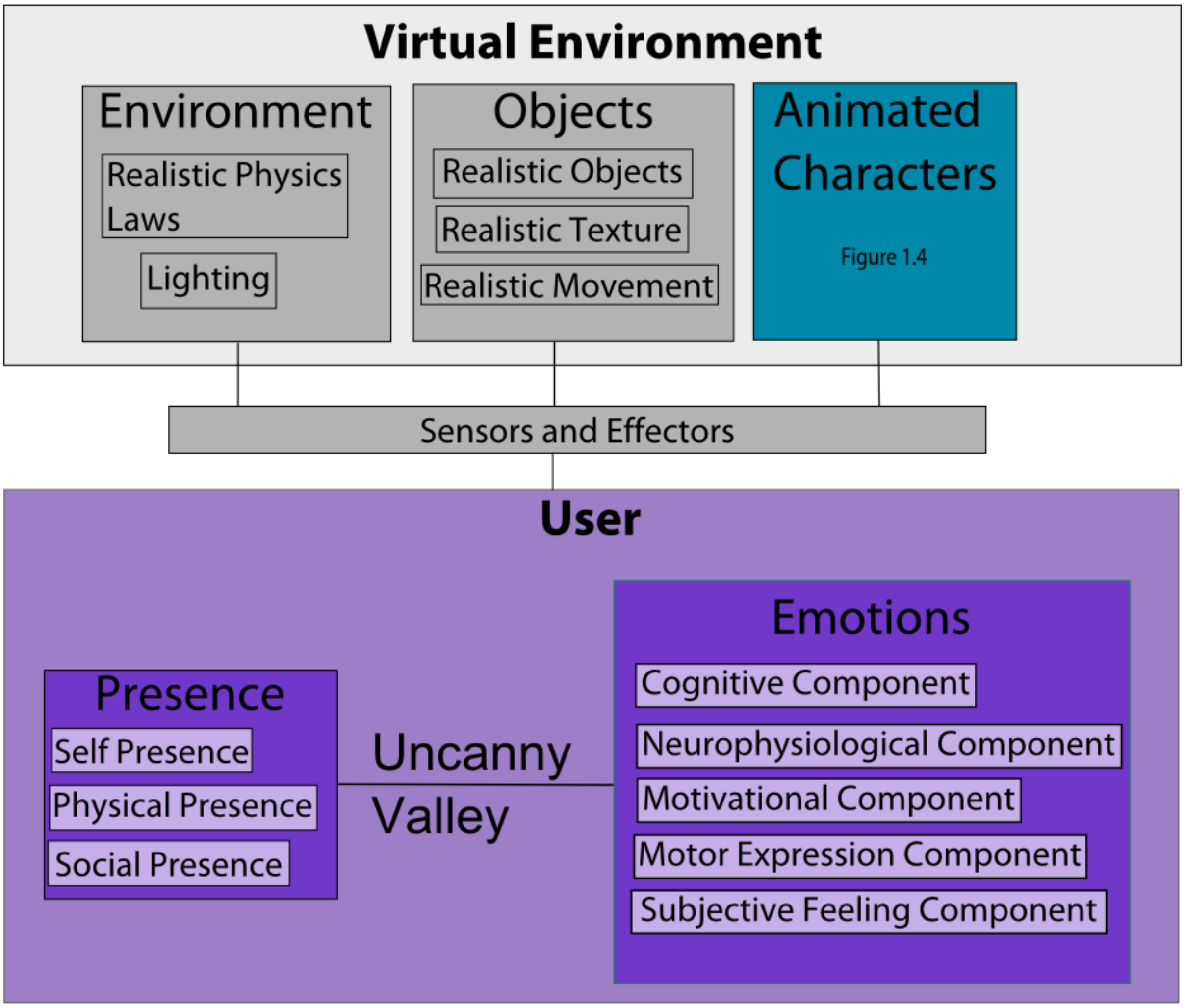


**Figure 1.8** Elements required for accurate transfer of skills lead to the Uncanny Valley.

The problem is that a non engaging character can lead to "breaks in [the] suspension of disbelief" or "breaks in presence" (Slater & Steed, 2000). A break in presence may be conceived of as an attentional shift away from the mediated environment and towards the physical environment, but with the possibility to still feel a sense of presence in the mediated environment, albeit to a lesser extent (Slater, 2003). In other words, the user stops being completely focused on the Virtual Environment and focuses back on the real world around them, i.e. their sense of 'self' is no longer 'there'. It has been argued that, in order to maintain presence, a Virtual Environment should be able to match users' expectations (Nunez, 2004) and that a break in presence arises when parts of the environment are not consistent with the rest. As the realism of a Virtual Environment increases so do users' expectations. It therefore suggests that highly realistic virtual environments are more likely to fail to induce presence, as they are more likely to fail to match all the user's expectations.

This is comparable with what the Uncanny Valley describes as a drop in believability as a result of characters becoming highly realistic. Both, the Uncanny Valley and the phenomenon described as “break in presence” may work in a comparable manner, as they both seem to depend on consistency in the level of realism that is presented, across all aspects, such as appearance and movements (Figure 1.5).

Although, the evidence for an Uncanny Valley is anecdotal, it threatens the implicit assumption in developing hyper-real characters; that viewers perceive and interpret a humans’ and animated characters’ expressions in a similar way. The model predicts dissimilarities as the animation reaches a certain but ‘unknown’ level of realism. If the model is correct, using Virtual Environments populated with realistic animated characters for social task training may not be suitable. Thus, although efforts are being put in to creating such Virtual Environments from a technical perspective, it is not known whether they could be used effectively. This is especially problematic where realism is a key aspect of the training, such as the perception of subtle emotional cues. Thus, this thesis investigates the assumption that the perception and interpretation of a ‘real’ human’s and an animated character’s emotional body language is similar.

# 1.6. Thesis Statement and Overview

## 1.6.1. The Research Question

The Uncanny Valley, at a conceptual level, raises the question of how realism affects the perception of animated characters (see Section 1.5).

Realism can be decomposed in terms of Behavioural Realism and in terms of Physical Realism (see Section 1.3). Behavioural realism is related to the movements displayed (Figure 1.5). It defines how similar to human movements the animated movements are. Thus, in the context of this thesis, it includes appropriate facial and body movements (Figure 1.5). To date, research on the expression of emotions has mainly focused on understanding facial and vocal expressions (Kleinsmith, De Silva, & Bianchi-Berthouze, 2006), emphasising their importance. Similarly, existing research on the Uncanny Valley has focused on facial expression (Brenton, et al., 2005; Hanson, 2006; MacDorman, et al., 2009). However, recent studies on the expression of emotions have also shown the importance of body language as a medium to express emotions (den Stock, Righart, & de Gelder, 2007; Kleinsmith, Bianchi-

Berthouze, & Steed, 2011; Kleinsmith, et al., 2006). This suggests that the Uncanny Valley could also be due, in part, to poor animated body language as well (i.e. not behaviourally accurate). This possibility has not been previously investigated, even though emotional body language is important for social task training. Thus, for this thesis, emotional body language was selected as the modality to investigate the effect of behavioural realism on the perception of emotion. However, the uncanny valley suggests (Section 1.5), this may interact with physical realism.

Physical Realism is related to the appearance of a character (Figure 1.5). It defines how similar to a human an animated character looks. Thus, it includes texture and physical shape (Figure 1.5). However, since emotional body language was selected as the modality, texture is a limited channel for displaying emotion as facial expressions would be removed (e.g. no blushing), and typically clothing would prevent visual display of other textural changes. Thus, in the context of this thesis, the basic form of Physical Realism is framed in terms of a body's anatomical fidelity, rather than textures (Figure 1.6).

The effects of realism and the possible interaction between physical realism and behavioural realism predicted by the Uncanny Valley could result in the viewers' inability to accurately interpret an animated character's body language and this could have many possible outcomes (Chapter 3 and Chapter 4). For example, it could result in misperception of the cues themselves, where an incomplete or erroneous set of cues could be seen, or the cues might not be experienced (perceived or interpreted) at the intended strength. The cues could be interpreted as resulting from a different cause; either a different emotion, or simply as the results of an algorithmic loop (i.e. not from a character but from a machine). However, it might also result in a viewer perceiving the emotions as unbelievable, i.e. not resulting from an inner emotional state (Figure 1.5). Finally, the animated movements may be perceived as unnatural (e.g. machine like movements), which ultimately makes the entire experience appear unbelievable.

Therefore, to investigate these possibilities, direct comparisons are made between the perception of human emotional body language and the perception of the same body language displayed by different animated characters (see studies reported in Chapter 3 and Chapter 4) in terms of interpretation, emotional strength, believability and naturalness (see Section 3.3).

This led to the following research question:

***Does behavioural and/or physical realism affect the interpretation, perceived strength, believability and naturalness of animated emotional body language?***

## 1.6.2. Contribution to Knowledge

The aim of this research programme is to assess the effect of physical and behavioural realism on the perception of emotional body language displayed by animated characters (Section 1.6.1). This has led to contributions to knowledge in several areas.

### 1.6.2.1. Contribution to Theory

The main finding of this thesis is that emotional body language displayed by animated characters can be identified as accurately as emotional body language displayed in a recording of a real person (see Section 6.2 for discussion). Thus, the physical realism (i.e. the way characters look) does not affect the accuracy of recognition of the body language being displayed.

However, it was found that both characters' physical realism and behavioural realism (see Section 1.6.1 for definitions) affect the perceived emotional strength, believability and naturalness of the body language being displayed (see Section 3.6.2 and Section 6.4). More precisely, it was found that realistic behaviour is more appropriate when displayed by a realistic character whereas stylised display is more appropriate when displayed by a simplified character. This suggests that the Uncanny Valley could be resulting from a mismatch between the way a character looks and the way it behaves. As characters look or behave realistically, the level of expectation is raised, which induces a drop in believability if there is a mismatch. In other words, the results suggest that consistency should be kept between a character's appearance (i.e. anatomy) and the way it moves (i.e. emotional body language) (see Section 6.5).

There was, however, another issue with the Uncanny Valley model. It was found that differences in emotional intelligence (in participants) were related to the perceived believability of a character (see results reported in Section 3.5.7 and Section 5.9) and hence to its position on the Uncanny Valley. This suggests that a simple one-to-one relationship between believability and realism is an over-simplification of the reality (see Section 6.7 for discussion). It would be interesting to investigate if other factors have an effect as this could help to resolve some of the controversy surrounding the possible causes of the Uncanny Valley (Dautenhahn & Hurford, 2006).

However, even without attributing the results to a higher theory of perception, the findings have significance for the animation of characters for social task training.

#### 1.6.2.2. Contribution to Practice

It was found that the level of anatomical realism does not affect the recognition (Accuracy) of an emotion displayed by body language (see Section 6.2). So, animated characters could, in theory, be effectively used in Virtual Environments for social task training to display realistic emotional body language that can be identified appropriately. However, even a realistic character behaving appropriately had reduced emotional strength, believability and naturalness when compared to a real human (see Section 6.3 and Section 6.4 for discussion). This could be an issue for social task training because the characters need to be believable in order to trigger social presence (see Section 1.2). However the limits in which this applies are also discussed (see Section 6.4 and Section 7.2).

#### 1.6.2.3. Tools and Technique

Given the lack of empirical work in this area, the procedures and materials used to explore the research question are also unique and provide a contribution to knowledge; in addition to the results (reported in Section 3.5 and Chapter 5). The work reported in this thesis shows that Motion Capture Technology can be successfully used to create material for the study of emotional body language (see results reported in Section 5.5.1). This makes it possible to empirically test the accuracy of perception of humans and animated characters behaving similarly. Moreover, this has practical advantage for the study of emotional body language as it is possible to create and control the way it is displayed and hence use animation to investigate specific research questions in a controlled environment.

In addition, in order to collect the empirical data, the Geneva Emotional Wheel (Scherer, 2005) was used. This is usually used for participants to report their own internal emotional state, but here has been adapted and used for participants to report on their interpretation of body language. The results show that this questionnaire can be used to record the perception of emotional body language, which extends its usefulness. This is important as it has implications for the investigation of the psychology of emotion.

#### 1.6.2.4. Contribution to the Psychology of Emotion

It was previously established that body language can be accurately identified independently from facial and vocal expressions. However, most of the studies used a limited set of emotions in combination with forced choice questionnaires offering a very limited set of answers (Atkinson, et al., 2004; den Stock, et al., 2007; Kleinsmith, et al., 2006). This could have resulted in participants discriminating between the different emotions rather than identifying them (Winters, 2005). Moreover, due to the limited set of emotions used in these studies, their results do not necessarily imply that it is possible to interpret a wide range of emotions. However, the Geneva Emotional Wheel which was used in two experimental studies (reported in Chapter 3 and Chapter 4) offered a wide choice of emotional labels (Figure 3.3), confirming that participants were able to recognize the different emotions presented.

Moreover, the material used for the studies reported in this thesis included a wider set of emotions and included social emotion such as shame and pride, something that has not been attempted previously with animated characters. Participants were able to correctly interpret them when displayed through body language only. Hence, the recognition rates obtained suggest that the expressivity of body language can be extended beyond the basic emotions as the participants were able to identify a set of ten emotions watching body language only (see results reported in Section 5.1).

Since the nature (computer generated or biological) of a character does not affect the recognition of an emotion, animated characters can be used effectively in psychology studies. The body language displayed by an animated character can be manipulated in a controlled way. This makes it possible to study the effect of very specific features. For instance, animated characters could be used to systematically study the effect of context on the interpretation of emotional body language.

## 1.7. Structure of the Document

In this chapter, the concept of Presence and its importance for the use of Virtual Environments for social task training was defined (see Section 1.2). More precisely, it was suggested that Presence is essential for the use of Virtual Environments for training as it is only through Presence that trainees would ‘feel’ and behave as if the situation was real. However, the concept of the “Uncanny Valley threatens Presence, and in this needs to be considered more fully (Chapter 2), along with the reasons why traditional animation has not been affected by this issue (Appendix A).

This is followed by a comparative study, which investigated if emotional body language would be an appropriate modality for animated characters to display emotion. This was accomplished by comparing the interpretation of emotional body language displayed by animated characters and real actors (see Chapter 3). The results are discussed in relation to the Uncanny Valley (see Section 2.2 for definition). The results showed that, animated body language can be accurately interpreted (see Section 3.5). However, the videos of the actor were found to be more emotional, more believable and more natural than the animated characters, whilst displaying the same emotional body language (see results reported in Section 3.5). Moreover, there was a significant difference in the number of correctly interpreted negative emotions displayed. Although, not for positive emotions. It was suggested that this could be due to the physical appearance of the animated character or to the loss of micro-gestures inherent to Motion Capture technology (see discussion in Section 3.6).

Thus, a comparative study was conducted to investigate the potential causes for this drop in believability and lower negative emotion recognition (see Chapter 3). It investigated the effect of changing the level of physical realism of the animation as well as deteriorating the quality of the emotional body language itself. Whilst no effect was found regarding the deterioration of the emotional body language, the results show that the real actor was found to be more emotional, more believable and more natural than the two animated characters. The results are discussed in relation to the Uncanny Valley (see Chapter 6).

The consequences of these findings for the use of Virtual Reality for social skills training are discussed (see Chapter 7). The thesis concludes by showing how the contribution to knowledge is already applied in a related domain (see Section 7.4), and discussing what future work still remains to be completed (see Section 7.6).

# Chapter 2 - Investigating the Uncanny Valley

## 2.1. Introduction

According to the Realism Maximization Theory (Groom et al., 2009), increasing realism should improve the interactions between trainees and animated characters. It should be noted that realism in this context includes anatomy (the way a character looks) and behavioural realism (the way a character moves) (see Section 1.6.1 for definition). Moreover, the highly realistic character needed for social task training should also increase the feeling of Presence (see Section 1.2 for definition). This theory is supported by empirical research showing that people are more comfortable interacting with those that look most like them (Koda, 1996; Wexelblat, 1998), suggesting they would prefer interacting with a realistic character, i.e. one that resembles a human being. Research has also shown that users prefer conversing with agents with faces, rather than agents without faces (Takeuchi & Nagao, 1993), and that agents with facial expressions are also preferred (Koda, 1996). Moreover, animated agents are preferred to still images (McBreen & Mervyn, 2001).

However, in most of these studies the level of realism has been tested by completely removing a parameter rather than varying its level, i.e. the agent's face (Takeuchi & Nagao, 1993) the agent's expressions (Koda, 1996), the agent' movements (McBreen & Mervyn, 2001). It is therefore not clear whether the effect found was due to a higher level of realism (either behavioural or anatomical), or simply to the inclusion (or absence) of this parameter (Figure 1.5). However, the situation may not be 'binary' and it can be suggested that the effect of realism is likely to depend on the 'level' that a character reaches, in other words on how similar to a human it seems.

Computer graphics have made huge progress in recent years and the level of realism that can now be created far exceeds the quality that was used in those studies. Due to technical limitations, the study supporting the realism maximization theory could not use "highly" realistic characters and hence did not investigate the effects of characters that are very similar to humans. Additionally, the Realism Maximization Theory has been strongly challenged and it cannot be simply assumed that greater fidelity would improve the interaction.

## 2.2. Definition and Context of the Uncanny Valley

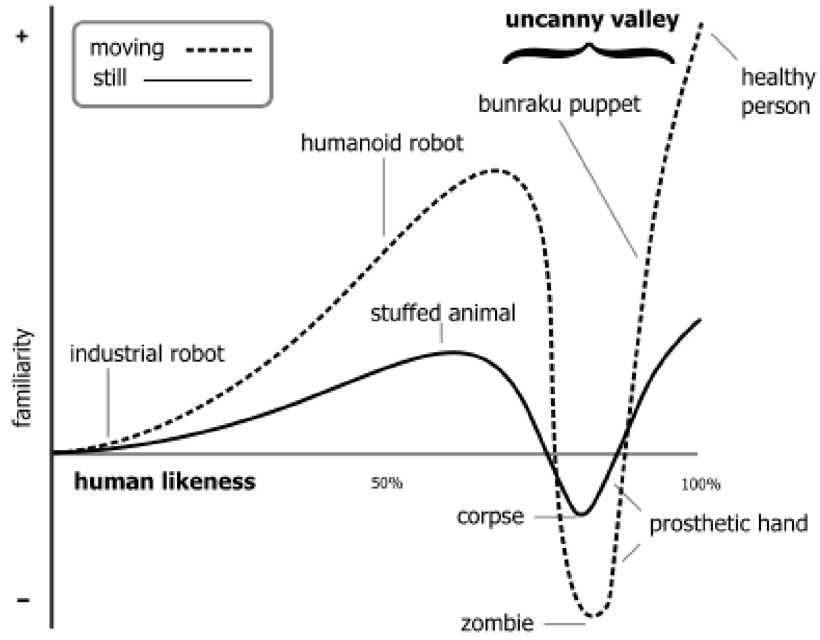


**Figure 2.1 :** The Uncanny Valley (adapted from Mori 1970) (Mori, 1970)

As computer generated or robotic agents become more visually realistic, they are confronted with the well-known Uncanny Valley problem (Mori, 1970). The Uncanny Valley (Figure 2.1) models a drop in believability as agents acquire greater visual similarity with humans (Brenton, et al., 2005). The concept was first introduced in robotics, where it was reported that highly realistic humanoid robots tend to be found repulsive (Mori, 1970). The axis labelled *human likeness* (Figure 2.1) describes the physical properties of the viewed 'target' and is a measure of its physical realism (how similar to a human it looks). The ordinate axis, labelled *familiarity* (Figure 2.1), describes not how real it looks but how believable it seems or 'feels'. Mori describes the uncomfortable feeling of shaking a highly realistic prosthetic hand as it looks like a real hand but does not "feel" natural to the touch, which renders the experience unbelievable (Mori, 1970). Thus, the Uncanny Valley may, according to Mori (1970), be the result of a mismatch between the perceived properties of an experienced object, such as its visual form, how it physically feels, or how it moves or how it reacts.

**Figure 2.2:** Screen shot from *Polar Express (Left) and from Final Fantasy (Right)*

In recent years, improvements in the field of animation technology has increased the level of visual realism that can be achieved for computer generated characters (Figure 2.2). In the context of animation, realism can be defined in term of behavioural realism (i.e. how similar to a human's behaviour is the animated character's behaviour) and in term of physical realism (i.e. how similar to a human does the animated character look) (see Section 1.6.1). The expectation was that as visual realism increases so should believability. However, animated characters from the film industry have also been confronted with the same drop in believability described by the Uncanny Valley. For example, the characters from "The Polar Express" (Zemeckis, 2004) or "Final Fantasy" (Sakaguchi & Sakakibara, 2001) have failed to be convincing (MacDorman, et al., 2009) (Figure 2.2). These characters look physically realistic, however there is an undefined element that does not seem natural, which reduces the believability of the whole experience.

Thus, in animation, the Uncanny Valley is in essence a decrease in believability as animated characters become more visually realistic, past a certain acceptance level, perhaps due to a discrepancy with other expected properties. This may be the reason why traditional animations (e.g. Disney, Pixar Studios) aim to create a visual style at the "realism peak", just before the valley (Figure 2.2) instead of creating photo-realistic characters.

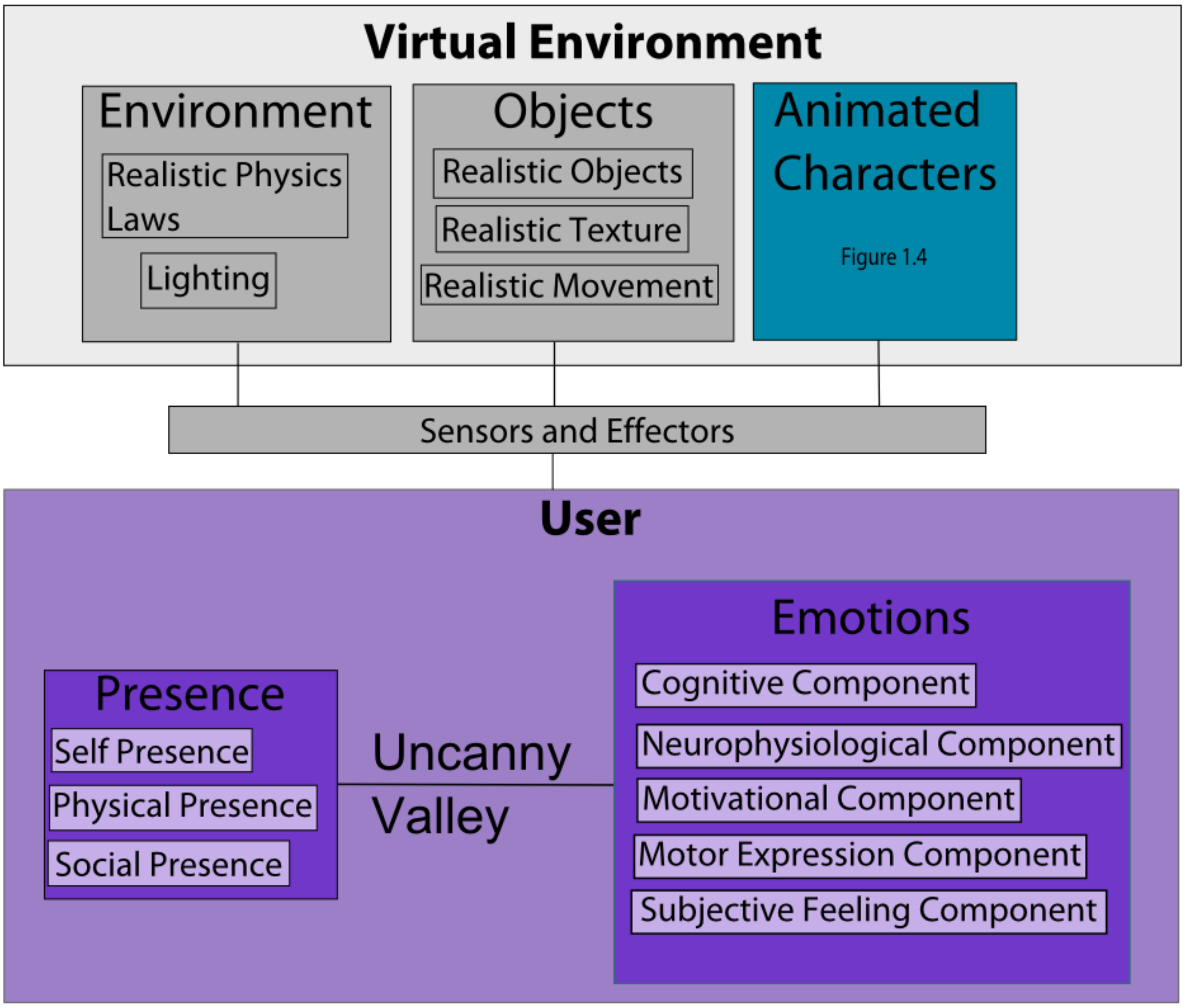


**Figure 2.3** The conditions necessary for accurate transfer of skills are potentially similar to those included in the Uncanny Valley.

For accurate transfer of skills to occur, it is essential to create situations comparable to real life. The problem is that by doing so, the characters needed for social task training may be perceived as uncanny (Figure 2.3). Uncanny characters (near “zombie” in Figure 2.1) that “appear real” but do not “feel real”, would not allow trainees to feel present in the Virtual Environment as it would constantly remind them that the situation was not real and therefore break their “suspension of disbelief” (Coleridge, 1817) (Figure 2.3). Unrealistic characters, such as the ones used in traditional cel animation, were not confronted with this issue as they produced an illusion of life (Appendix A) but would be problematic for accurate transfer of skills.

Although widely supported with anecdotal commentary, Mori's model is the product of his intuition. The Uncanny Valley was not grounded on systematic studies and its existence is thus subject to some debate elsewhere (Dautenhahn & Hurford, 2006).

# 2.3. Empirical Evaluations of the Uncanny Valley

## 2.3.1. Research supporting the existence of the Uncanny Valley

**Figure 2.4:** Example of faces used by MacDorman and Ishiguro (2006)

An attempt to empirically test the Uncanny Valley was made in 2006 by MacDorman and Ishiguro. In order to investigate the shape of the Uncanny Valley they used static images of a non-realistic robot that was morphed at different levels with a realistic humanoid robot that in turn morphed at different levels with a picture of a human (Figure 2.4). This resulted in pictures of faces that vary in their level of physical realism from very low realism (left of Figure 2.4) to completely human (Right of Figure 2.4). The experiment confirmed the existence of an Uncanny Valley as it was found that a high level of realism seems to result in an eerie feeling whereas both the human and the non-realistic humanoid faces were not rated as eerie (MacDorman & Ishiguro, 2006).

However, a confound could be present in the way the material was created. The pictures used in the study were automatically generated and present 'strange' features. For instance, the low realistic robot had a camera on the forehead, which did not completely disappear when morphed with more realistic images (Middle of Figure 2.4). Rather than the increase in physical realism it could be the blend of biological and synthetic features (i.e. incorrect anatomy) that is responsible for the eeriness found in the study. This is consistent with Seyama and Nagayama's (2007) findings, which used similar stimuli. Their results confirmed the presence of an Uncanny Valley when characters become more realistic, however, this was only the case for characters with abnormal features (Seyama & Nagayama, 2007). In a recent study, MacDorman et al.(2009) refined their methodology and assessed computer generated faces that vary in terms of facial proportions, skin texture, and level of detail. This study found that the most negative reaction were for faces where the photorealism of the eyes and the photorealism of the skin (i.e. incorrect textures) were mismatched suggesting that an Uncanny Valley exists and that it could result from discrepancies in levels of physical realism between aspects of a character (MacDorman, et al., 2009).

### 2.3.2. Research contradicting the existence of the Uncanny Valley

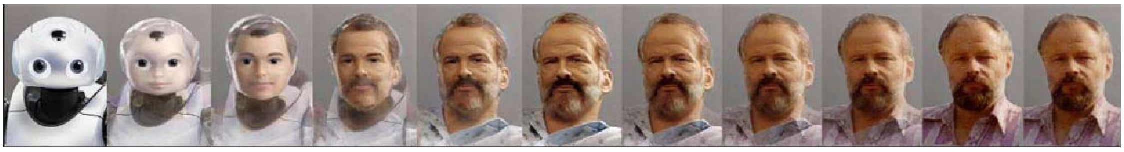

**Figure 2.5:** Example of faces used by Hanson (2006)

However, Hanson (2006) stated that a "well designed" character will be accepted independently from its level of realism. Hanson (2006) argued against the Uncanny Valley and describes the problem as an aesthetic one. According to his theory, the acceptance of a robot or an animated character depends on aesthetic features (Hanson, 2006). However, no formal definition of 'aesthetic' is proposed by Hanson (2006), leaving his theory unclear. In an experiment, using the same experimental procedure as MacDorman & Ishiguro (2006) study, Hanson found supporting results for his theory. The difference between the two experiments was on the method used to construct the tested material. The pictures used for this study were not automatically morphed together but were morphed to be similar in terms of realism like the ones used by MacDorman but differed in term of aesthetic (Figure 2.5). By carefully constructing the different pictures and making sure that a good aesthetic was maintained, the Uncanny Valley depicted in MacDorman work was not found (Hanson, Olney, Pereira, & Zielke, 2005). These results challenge the nature of the Uncanny Valley as they suggest that it could result from poor design only. However, looking at the material used (Figure 2.5), the level of realism does not seem to increase regularly from picture to picture. This could also explain why no evidence of the Uncanny Valley was found. Moreover, these results are not consistent with the examples from the movie industry that tends to favour an appealing aesthetic design but also created some characters that were generally poorly accepted (see Section 1.5 for examples).

## 2.4. Potential causes of the Uncanny Valley

**Figure 2.6** : Kaspar, a Humanoid Robot for Human-Robot Interaction Research

The robot Kaspar, a child-sized humanoid robot (Dautenhahn et al., 2009), is anecdotally described as uncanny or scary (Figure 2.6), suggesting that it falls in the Uncanny Valley. However, its "uncanny appearance" does not stop users from interacting and engaging with it. The robot has been found to be socially engaging and proven successful in evaluation studies (Figure 2.6) (Robins & Dautenhahn, 2007), hence overcoming the issue raised by the Uncanny Valley. However, this success could be due to the population sampled as the robot was evaluated with children. This is supported by findings from Ho (2008) who reported that women were found more sensitive than men to the phenomenon (Ho, MacDorman, & Pramono, 2008). Taken together, it suggests that individual differences, such as gender, age etc., might affect the Uncanny Valley. Kaspar's success could also indicate that the Uncanny effect fades over time as users get used to the appearance of a character. Thus, the existing studies do not seem to fully explore the full complexity of the problem, which may involve a complex combination of individual, contextual, cultural and social factors among others (Cañamero, 2006).

Brenton and colleagues (2005) suggest that the Uncanny Valley result from poor facial animation. However, traditional animation that creates a wide range of credible characters, avoiding the Uncanny Valley, highlights the importance of displaying emotional body language (Bates, 1994; Thomas & Johnston, 1995). Research on the expression of emotions has mainly focused on understanding facial and vocal expressions (Kleinsmith, et al., 2006), emphasising their importance. However, recent studies have also shown the importance of body language as a medium to express emotions (den Stock, et al., 2007). This suggests that the Uncanny Valley may also be due to poorly animated body language as well. Thus, an animated character displaying emotion realistically through the face and not through the rest of the body may look as unnatural to a viewer as a character with no emotional expression at all. This could be due to poorly animated movements or to discrepancies between the facial expressions and the body movements. The possibility of body expressions causing the Uncanny Valley remains unexplored as most of the studies on the Uncanny Valley have focused on facial expressions, see for instance (MacDorman, et al., 2009) (Hanson, et al., 2005).

MacDorman (2006) has suggested that the Uncanny Valley could result from poorly animated facial expressions. This is supported by brain imaging studies that showed that the recognition of facial expressions is anatomically and functionally specialized (Carmel & Bentin, 2002; MacDorman & Ishiguro, 2006). Those specialized mechanisms would be triggered while seeing imperfect animated faces, which in turn trigger an uncomfortable feeling. In other words, it would be the combination of imperfect facial animation and of the brain's extra sensitivity to facial expressions that would trigger the negative reactions described by the Uncanny Valley.

However, it has also been established that the recognition of body expression triggers specialized brain mechanisms, which overlap with the ones dedicated to facial expression recognition (Van de Riet, Grèzes , & de Gelder, 2009). Thus, if the Uncanny Valley results from specialized brain mechanisms, these results suggest that body expressions would be affected as well. Similarly, it has been suggested that the Uncanny Valley could result from brain mechanisms that are both active when watching someone perform an action and when performing the same action (i.e. the mirror neuron) (Rizzolatti & Fabbri-Destro, 2009). It has been suggested that mirror neurons are activated while watching 3D animated characters as well (Chaminade, Hodgins, & Kawato, 2007) and viewers could be highly sensitive to imperfections, that could also result in the uncomfortable feeling described by the Uncanny Valley.

However, the Uncanny Valley could also result from a threat avoidance mechanism (Mori 1970). A poorly animated character could trigger the same reactions as seeing someone carrying a potentially contagious disease. This would explain why the Uncanny Valley applies to highly realistic characters, as a human would be endangered by disease carried by organisms closely genetically related to them (MacDorman & Ishiguro, 2006). Similarly, it has been suggested that realistic characters with imperfections could, to a certain extent, remind viewers of their own mortality and would therefore induce these negative reactions.

However, such mechanisms would not apply to poor facial animation only. For instance, a poorly animated walking cycle or an unnatural body posture could trigger similar avoidance mechanisms as human could be sensitive to poor body animation as well as poor facial behaviour. Thus, the rationales that have been proposed to explain the Uncanny Valley remain to be validated and could apply to whole body animated behaviour as well as animated facial behaviour. Further discussions on the nature of the Uncanny Valley can also be found in (Dautenhahn & Hurford, 2006).

## 2.5. Conclusion

The relationship described by the Uncanny Valley between physical realism and believability seems comparable to the relationship between realism and presence in a virtual environment. In a Virtual Environment, it is difficult to maintain presence and a high level of realism simultaneously. This suggests that both phenomena are better explained by a mismatch between the way the character, or the environment, looks and the discrepancy between users' expectations because of this and the way it behaves or reacts to users' actions. This is coherent with the consistency theory (Groom, et al., 2009), which argues that realism does not necessarily dictate users' responses, but it is maintaining consistency that avoids the Uncanny Valley. Furthermore, this theory is not necessarily in opposition with the Uncanny Valley as consistency becomes harder to maintain as realism increases and could therefore explain the drop described when realism is increased beyond a certain level. This theory would also explain the results of MacDorman (2009) and Hanson (2006) studies. The drop in believability that was found in MacDorman's study could be due to the inconsistencies created by the morphing pictures (see Section 2.3). In Hanson's study, this drop in believability was not found once the inconsistencies created by the automatic morphing were removed. This is also supported by Nowak & Biocca's (2003) findings, which reported that users felt more presence when a low anthropomorphic representation was used rather than a high anthropomorphic one (Nowak & Biocca, 2003). The authors suggest that this difference was due to the rise in users' expectations when interacting with a realistic representation.

Although it has long been accepted in traditional animation, that expressions must be captured throughout the whole body as well as in the face (Thomas & Johnston, 1995), most of the studies considering the Uncanny Valley used static images of faces. Thus, it is not evident that realistic emotional body language displayed by an animated character, rather than by a static character, and by humans is similarly interpreted and has the same effect. In other words, there is no evidence that the perception of emotions displayed by an animated character can be accurate. This would obviously raise serious issues regarding their usage as a training tool for areas such as medical interviews.

Motion capture has been used to accurately capture an actor's gestures and hence animate a character's whole performance (Section 2.2). Indeed, Andy Serkis was nominated for an award for his motion captured performance in Lord of the Rings (Jackson, 2001). In these animations, body movements and gestures are never displayed independently of the face and hence it is unclear whether the uncanny effect is really due to poor facial animation or a combination of that and other factors, such as unrealistic muscles deformation (i.e. no secondary animation). Indeed, consistency theory (Groom et al., 2009) suggests that the mismatch may be the problem. Motion Captured data can however be used to animate emotional body language in isolation, and hence make it possible to investigate whether body language produces inconsistencies between displayed body language and perceivable body animation that does not meet viewer's expectations and hence contributes to the uncanny effect of hyper-realistic characters.

Thus, considering these issues and the importance of displaying emotions for animated characters (see Section 1.4), research will focus on emotional body language and will investigate whether realism affects the perception of emotional body language. This is important as recent results suggest that the Uncanny Valley may in fact, be a "cliff" (Bartneck, Kanda, Ishiguro, & Hagita, 2007), and that the problem is essentially insurmountable. In order to investigate this issue, it is necessary to investigate whether realism has an effect on the perception of an emotion displayed by an animated character. Realism can be defined in term of behavioural realism (i.e. how similar to humans' behaviour is the animated character's behaviour) or in terms of physical realism (i.e. how similar to a human does the animated character look).

Thus, although effort is being expended to create expressive agents from a technical perspective, it is not known whether they could be used in the context of natural interactions with humans as the perception (including believability) of realistic displays of emotion has not been investigated and needs to be for social task training.

# Chapter 3 - Body Language as a Way for Animated Characters to Display Emotions

## 3.1. Introduction

Computer based simulation has demonstrated effectiveness for medical training (Anolli, et al., 2006). These types of simulation have used emotional animated characters. Even though emotions have a strong influence on human-human interactions (Gratch, Mao, & Marsella, 2006), typical system evaluation does not assess whether human and animated emotional displays are perceived similarly by observers. In social task training, including medical interview, such animated characters would have to be perceived as social and emotional partners so that trainees would be confronted with situations comparable to real life ones in a Virtual Environment inducing presence.

However, due to the problem described as the Uncanny Valley, the appropriateness of the perception of a realistic emotional animated character needs to be investigated (see further discussion on the Uncanny Valley in Chapter 2). Using a similar approach as the one proposed by Nass & Moon (2000) in their work on the Media Equation, this chapter reports on a comparative study, which investigated how emotional body language displayed by an animated character and a human performer are perceived.

## 3.2. The Media Equation

Research has investigated how humans interact with technology from a social perspective, ultimately leading to the Media Equation (Reeves & Nass, 1996). Reeves & Nass used existing results from experimental social psychology that defined rules that apply to Human-Human interaction (Reeves & Nass, 1996) and investigated whether these rules apply to Human-Computer interaction as well. Their work showed that the way humans interact with technology is 'mindlessly' social. They found technology can trigger social scripts, which typically apply to human-human interaction but are inappropriate for human-computer interaction, as they ignore the essential nature of the technology (Nass & Moon, 2000). The social rules tested include perceived expertise, as media content tends to be rated more favourably when displayed via technology labelled as 'specialist'. Similar experiments demonstrated that computers can be considered as team-mates, prompting the activation of social rules governing such a relationship. Another striking example is politeness towards computers, where *"Adults have been shown to apply the same social norms and rules of etiquette toward computers as they do toward other humans"* (Aharoni & Fridlund, 2006). Consistent with the Media Equation, other studies support the fact that animated characters are indeed perceived as social agents and trigger natural and social protocols of human users, such as gaze (Prendinger, Ma, & Ishizuka, 2007) or different acceptance levels of animated character's recommendations based on ethnicity of the character (Pratt, Hauser, Ugray, & Patterson, 2007).

| | **Mindlessly perceived as social partner** | **Consciously perceived as social partner** |
|---|---|---|
| **Realistic Characters** | YES (Prendinger, et al., 2007) (Pratt A, et al., 2007) | ? |
| **Restricted Technology** | YES (Nass & Moon, 2000) (Aharoni & Fridlund, 2006) | NO (Nass & Moon, 2000) |

**Table 3.1**: Summary of the Media Equation results.

These experiments support the Media Equation even within very restricted technologies (Table 3.1) and show that users exhibit what the authors called mindless social reactions. However, when participants are asked whether their reactions towards the technology were social their answers were constantly negative (Nass & Moon, 2000). Therefore, "The Media Equation" did not remain true on a general basis as media is not consciously considered as a social partner (Nass & Moon, 2000).

However, this was with restricted technology and it may not hold true for highly realistic characters, as they can use all the features used in face-to-face interaction (Cassell, 2000). There is contradicting evidence regarding whether realistic characters are considered as social partners. For instance, Bickmore et al.(2005) reported successful interaction and high user engagement with their artificial agent Rea, which was designed as a virtual tutor for a patient to learn healthy behaviours. On the other hand, Raij et al.(2007) reported insincere engagement from users while interacting with their virtual patient. It should be noted though that the issue described by the Uncanny Valley also threatens the interaction and the engagement with realistic characters (see Chapter 2 for discussions).

Thus, for social task training, it is necessary to investigate whether an animated character would be consciously perceived as a social partner (Figure 3.1), and hence test the implicit assumption in developing hyper-real characters that viewers perceive and interpret humans and animated characters in a similar way. Moreover, a similar approach, as the one used to test the media equation, can be adapted to test the ability to perceive (i.e. see) and interpret (i.e. attribute meaning) emotional expressions when displayed through the body. Emotional body language is an ideal start to an investigation as it is known that people can accurately distinguish among emotions when displayed through human body language (den Stock, et al., 2007) and it is possible to create highly realistic animation using motion capture technology (Section 1.4).

## 3.3. Extending the Media Equation to Emotion

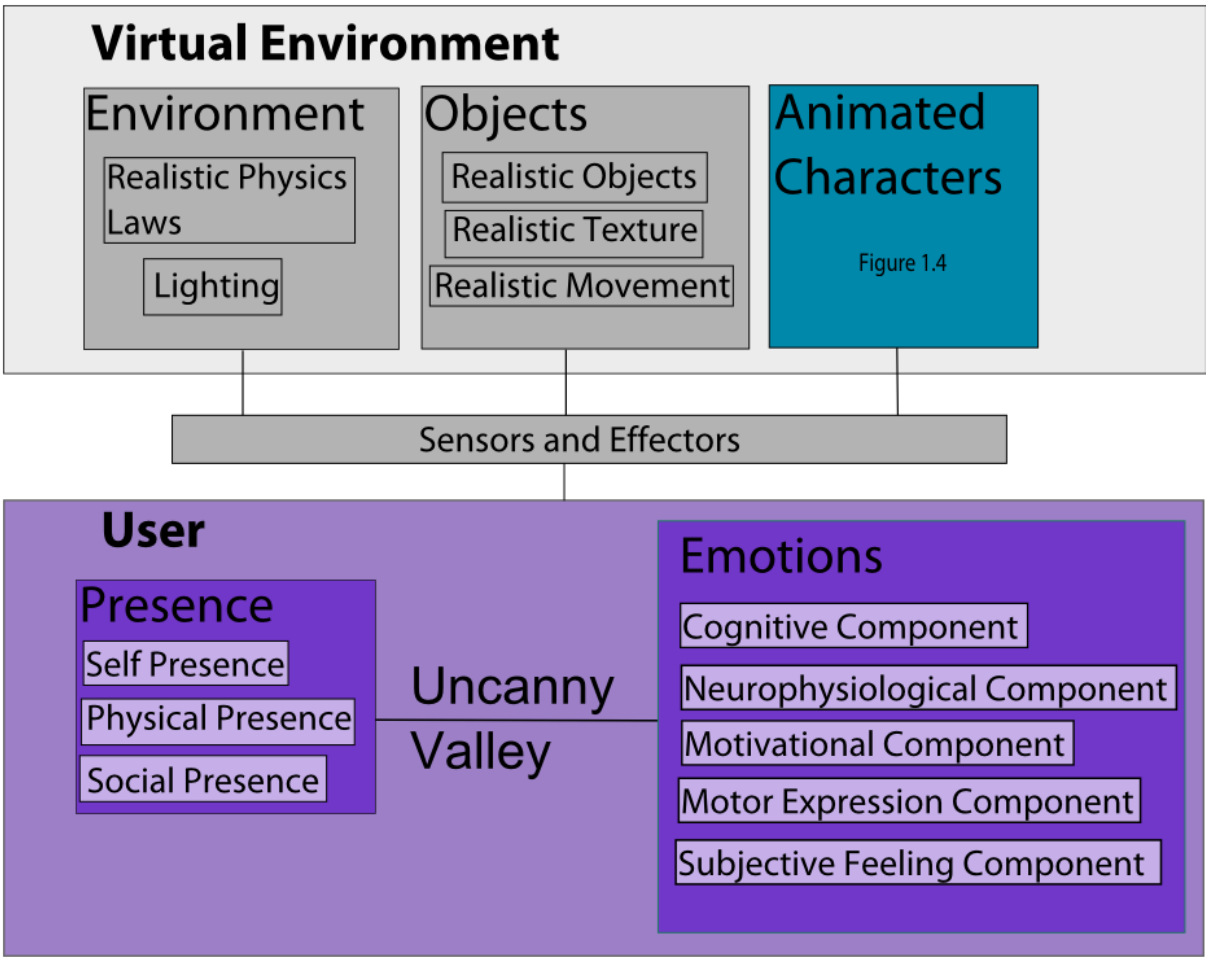


**Figure 3.1** Emotional Perception in the context of Virtual Environments

Exploring if the media equation held true at a conscious level for animated characters (important for training in tasks that require decision making) can be done by testing if viewers would perceive and interpret a human's or animated character's body language in a similar way. Testing this assumption can provide an avenue for an investigation into the Uncanny Valley and it is important to test whether or not the media equation holds true at a conscious level for animated characters.

In the context of this study, it can be argued that users' perception of emotion occurs within the Cognitive Component of emotions (Figure 3.1). Moreover, in the psychology of emotion, perception of emotion is often experimentally explored in terms of accurately identifying discrete emotional categories along with their intensity (Kalat & Shiota, 2007). This paradigm was used in order to make direct comparison between the ability to interpret emotional expression displayed by a human and an animated character put in the same context.

The study was designed to test the implicit assumption that viewers would perceive and interpret humans and animated characters in a similar way (i.e. would attribute the same label to the emotional body language displayed by a human and by an animated character). Differences in interpretation could be due to the Uncanny Valley, which is a drop in believability as animated characters become more realistic or to poorly animated movements (see Chapter 2 for discussions). Thus, conscious perception needs to be established in terms of accuracy, but also, in term of believability and naturalness of movements (Figure 3.1). If realistic animated displays of emotion are not correctly interpreted or perceived as unbelievable or unnatural, it may result in not developing skills that are transferable to real world situations.

The assumption that viewers would perceive and interpret a human's or animated character's body language in a similar way will therefore be investigated using a set of emotional performances recorded using Motion Capture Technology (motion capture technology accurately capture an actor's gestures which can then be used to animate a character's whole performance). This will then be used to compare viewers' perception and interpretation of the same emotional body language displayed either by a human or by an animated character.

Testing this assumption can provide an avenue for an investigation into the Uncanny Valley and it is important to test whether or not the media equation holds true at a conscious level for animated characters.

# 3.4. The Experiment

## 3.4.1. Design

The aim of the experiment was to compare emotional interpretation (*Correct Identification* and perceived emotional *Strength*) (Figure 3.1) of a recorded person and an animation (*Actor Type*), which were both displaying the same emotional body language (Actor Vs Animated). Thus, a formal experimental procedure was adopted, however, several issues in perception of animation have been identified that need to be controlled.

In order to avoid the Uncanny Valley (see Chapter 2), traditional animation usually uses stylised displays of emotion, which are adapted to the physical appearance of the character (Appendix A). This could affect the perception of the animated display (Participants on seeing an animated character could expect stylised and exaggerated display of emotions). Therefore the study also investigated whether there is an 'uncanny effect' due to the style of displays, by comparing the perceived *Believability* and *Naturalness*, of stylised and ordinary displays of emotion (*Action Style*).

Previous studies on emotional body language recognition (Atkinson, et al., 2004; den Stock, Grèzes, & de Gelder, 2008; Pollick, Paterson, Bruderlin, & Sanford, 2001), have considered small sets of prototypical emotions (i.e. happiness, sadness, fear, anger, disgust, surprise), which could have resulted in discrimination between emotions (a process of elimination) rather than true recognition. Therefore, a wider range of emotions was used for this study, including anger, disgust, shame, fear, sadness, positive surprise, relief, happiness, pride and excitement. This point is supported by Kramer (2007), who further argued that nonverbal behaviours rarely carry specific intrinsic meanings but are interpreted within context (Kramer, 2007). Therefore, this study also investigates the effect of body language in two different situations, one with contextual verbal support and one unsupported (Voice and Body Language vs. Body Language) on the interpretation of the Actor and Animated display.

Moreover, it should be noted that the set of emotions tested included two social emotions, pride and shame. They were specifically chosen to investigate whether they can be correctly interpreted when displayed by an animated character. This type of emotion may play an important role for social task training.

To further explore the Uncanny Valley, the study investigated whether there is a relationship between perceived *physical realism* of the characters presented and the *Believability* of the display. This variable was also used to explore the appropriateness of the *Action Style* for each character.

| Character Type / Modality | Actor | Animated Character |
|---|---|---|
| Voice and Body Language | Ordinary / Stylised | Ordinary / Stylised |
| Body Language | Ordinary / Stylised | Ordinary / Stylised |

**Table 3.2**. Experiment conditions. The effect of *Character Type* was tested within subjects. The effect of *Modality* was tested between subjects.

The experiment used a mixed design with three independent variables (*Character Type*, *Modality* and *Action Style*) (Table 3.2).

Four dependent variables were defined to explore the perception of the emotional body language displayed, *Correct Identification*, *Strength*, *Believability* and *Naturalness* (Figure 3.1). *Correct Identification* was used to test whether it was possible for participants to interpret the emotions displayed. *Strength* was used to analyse the perceived intensity of the emotions being displayed. *Believability* and *Naturalness* were both used to test the Uncanny Valley as a model. The Uncanny Valley predicts a drop in *Believability* as realism increases whereas *Naturalness* tests whether movements recorded by the motion capture were perceived as physically realistic.

Moreover, for this study, it was necessary to consider the participants' ability to accurately interpret emotional behaviour, displayed by the actor and the animated character, and how this may affect the results. Perception of emotions is reported as being one of the main components of Emotional Intelligence (Kalat & Shiota, 2007). Therefore Emotional Intelligence, including whether it correlates with the viewer's ability to classify emotion in either the video or the animated display, was investigated. Moreover, previous experience in video games and animated characters could also affect the perception of the animated character. Therefore, the experiments investigated possible correlations between previous experience and the results.

### 3.4.2. Participants

25 Participants were recruited, mostly members of staff of the University of Portsmouth (16 females and 9 males) ranging in age from 24 to 61. Participants were randomly allocated to one of the two groups (Voice vs. No Voice). The thirteen participants in the Voice condition ranged in age from 24 to 61 (M=39.3, SD=13.7). The twelve participants in the no voice condition ranged in age from 31 to 60 (M=44.9, SD=9.7). Participants were entered in a raffle in exchange for participation.

### 3.4.3. Apparatus

Before the initial pilot of the study, a professional actor was hired to record the performances. During this session, the actor sought direction from the researchers. However, the lack of professional guidance resulted in performances that were highly prototypical. During a debrief session, the actor explained the type of direction that was needed to achieve more natural results. Thus, a professional actor as well as a professional director were subsequently hired to perform the recorded performances, and a modified procedure was created.

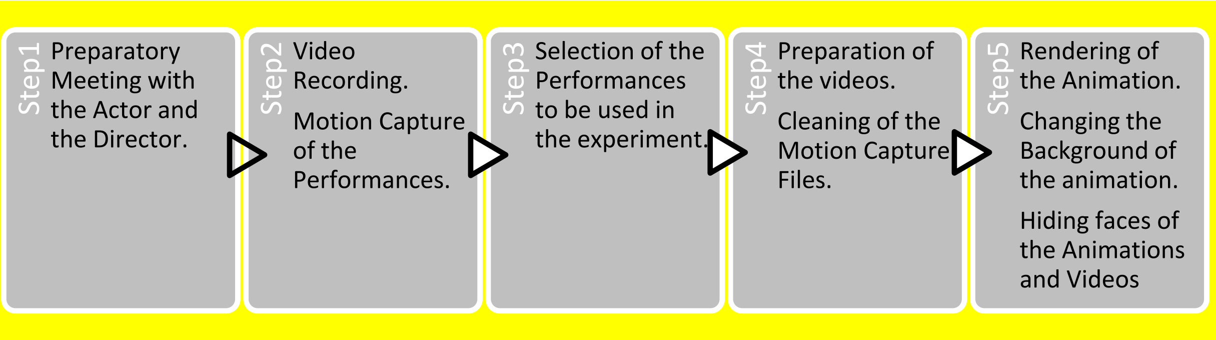


**Figure 3.2.** Steps to create the experiment's material

Prior to the recording, both the actor and director were briefed regarding the experiment (Figure 3.2). The emotions to be performed were chosen to cover different areas of the Geneva Emotion Wheel but also based on their ease to be performed. The selected emotions were: Anger, disgust, shame, fear, sadness, positive surprise, relief, happiness, pride and excitement. Each emotion was to be performed in two different ways, an ordinary version and an exaggerated stylised one, similar to that seen in traditional animation.

The director and the actor prepared two different set of sentences (In English) that were used during the recordings. These sentences did not contain any emotional information and were used during the recordings for all the emotions as well as for the two styles. The first one was a monologue by the actor:

(A) Hey you.

(A) hey you over there wait.

(A) hey you over there wait for me.

The second was a dialogue between the actor and the director:

(A) Why are you doing this?

(D) Doing what?

(A) This.

(D) I don't know what you mean.

(A) You know exactly what I mean.

During the recording, the actor followed the director's instructions. He continued performing each emotion until the director was satisfied with the performance. This ensured that the recorded display conveyed the intended emotion as well as being appropriate for the ordinary or stylised condition.

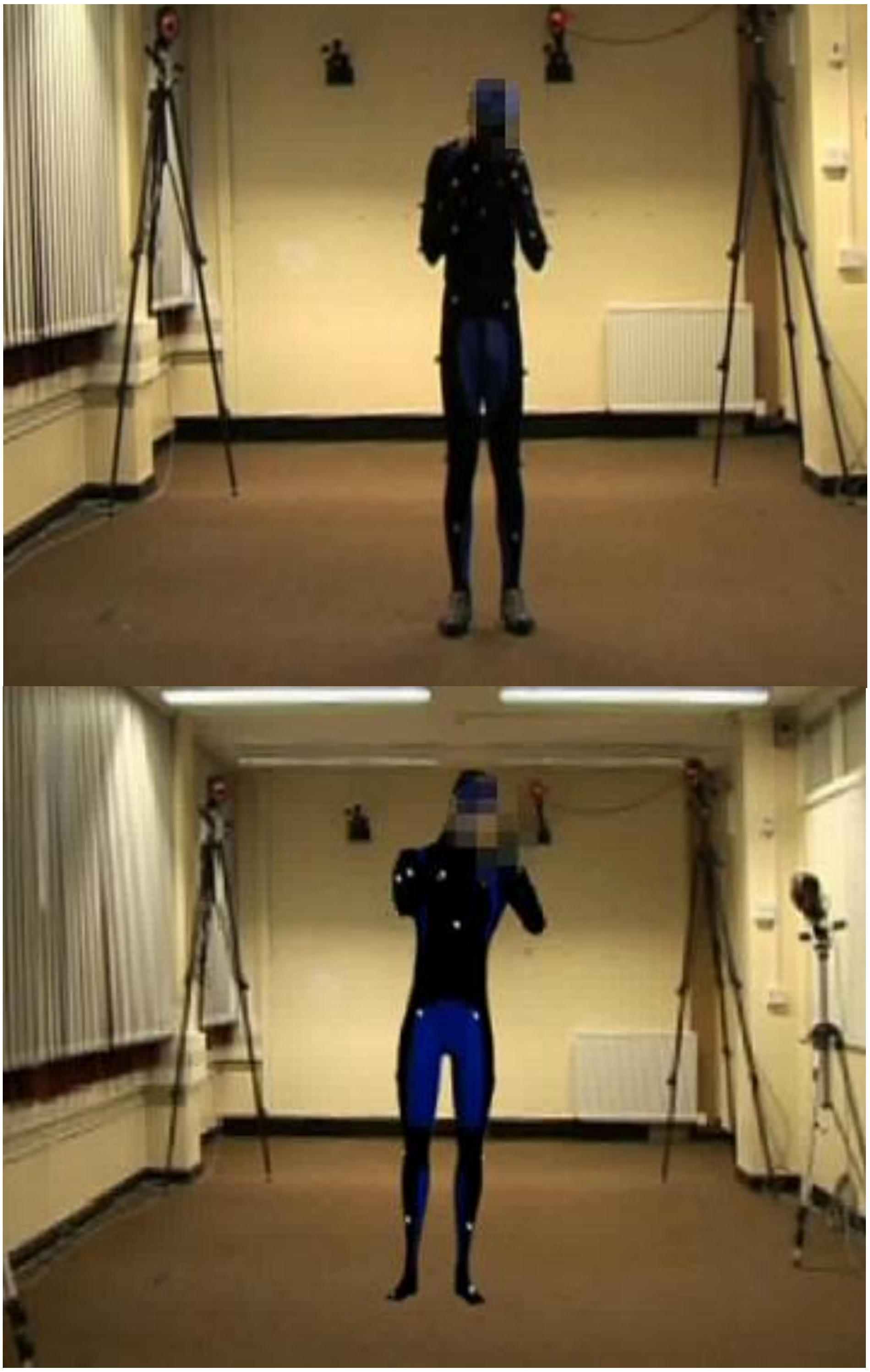

**Figure 3.3.** Screen shots of video condition (top) and animated condition (bottom).

To ensure equivalency across conditions, the actor was video recorded (Figure 3.3 top) and motion captured simultaneously. The videos were recorded using a Sony PD170P, at Standard Definition (525i in 4:3 mode). Motion capture data was recorded using a VICON optical motion capture system (Figure 3.3) (8 "v-series" cameras at VGA, 640x480, resolution, resolved with VICON IQ2.). The motion capture data was visually checked and cleaned when necessary and then used to animate a character (Figure 3.3 bottom) so that it displayed the same emotional body language.

The animated character used was a high polygon mesh. Autodesk 3DS Max was used to rig the character and to render the animation. When rendering the animation, a green background, equivalent to a chroma key, was placed behind the animated character. It was then overlaid with a video recording of the room used to record the motion capture using Adobe After Effects. The faces of the actor and the animated character were pixelated (using Adobe After Effects as well), so that this source of information was removed along with the possible uncanny effect that may come from facial animations (Figure 3.3). Due to the complexity of the lighting, almost no shadows were present in the actor condition (Figure 3.3). No shadows were rendered in the animated condition. Moreover, to remove possible effects, such as differences in dress of the actor and animated character, both appeared in a motion capture suit (Figure 3.3), were physically similar (skin and face are not visible) and were put in the same context (Figure 3.3). Forty 'presented emotion clips' were created: 10 emotions x 2 versions x 2 types of display (see table 2).

These presented emotion clips and were displayed on a 5m x 2.5m rear projection screen at life size. To record participants' answers, the material was embedded into custom made software, which was used for displaying the clips as well as recording participants' answers.

### 3.4.4. Materials

The *Emotional Intelligence* was measured using the Emotional Intelligence Test-2$^{nd}$ Revision (Jerabek, 2001).

The *Physical Realism* of the actor and animated character were rated once on a seven point Lickert scale using static pictures showing them in the same 'neutral' position. For each emotional clip, participants were asked to rate the *Believability* and *Naturalness* on a seven point Lickert scale. These were presented in the proprietary application.

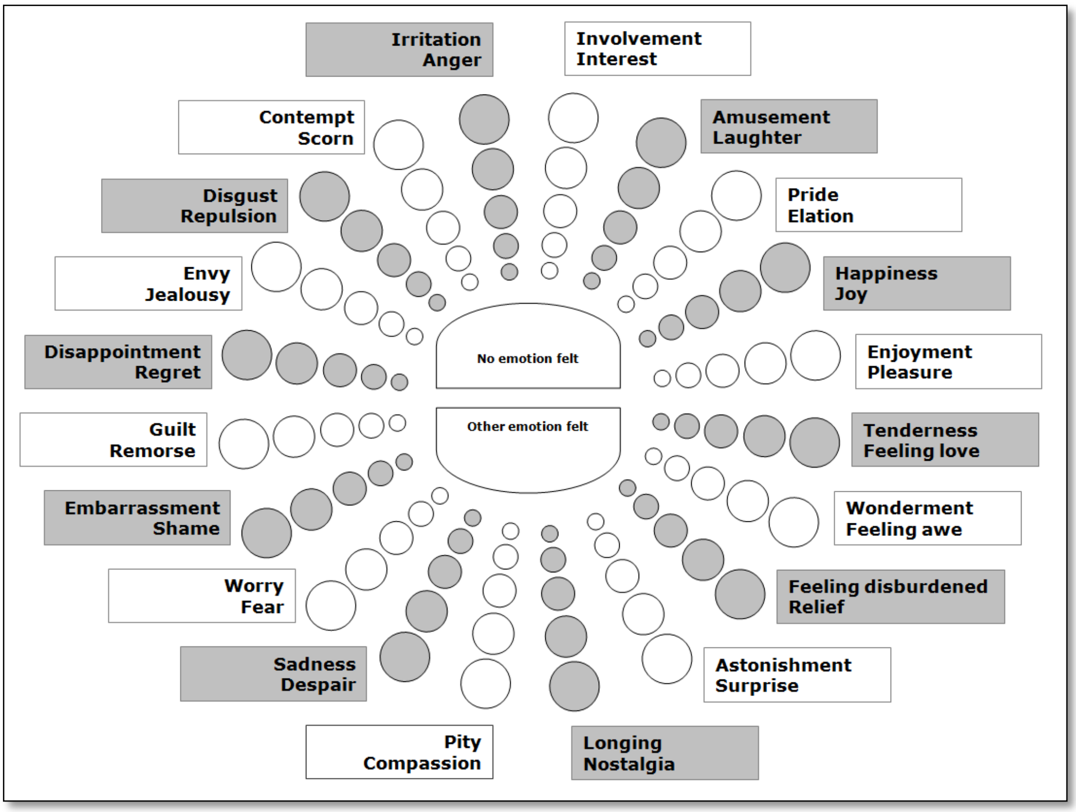


**Figure 3.4.** The Geneva Emotional Wheel (Scherer, 2005).

In order to record participants' interpretations of the emotional body language displayed, an existing questionnaire has been modified. It is based on the Geneva Emotional Wheel that places twenty emotional labels on a two dimensional axis: valence and control (Figure 3.4) (Scherer, 2005). The Geneva Emotional Wheel is usually used for self-report (i.e. participants reporting their own emotional state). However, in this study, it has been used to report on participants' interpretations of the clips. The wheel's centre includes two additional options, "no emotion at all" and "none of the above", in case an emotion is perceived by a participant that is not present on the wheel. Using this questionnaire, each participant was also able to report the strength for each presented emotion clip (five point Lickert scale radiating out from the centre) (Figure 3.4).

A post-study questionnaire gathered personal information including gender, age, experience of video games and familiarity with animated characters. Finally participants were asked whether they noticed the presence of an animated character in the presented emotion clips and if they found that character 'disturbing'.

### 3.4.5. Procedure

Participants were tested by the same experimenter but in individual sessions. Each session began by obtaining consent, followed by the emotional intelligence questionnaire.

After completion of the questionnaire, participants rated the realism (*Physical Realism*) of the actor and the animated character watching static pictures. It was necessary to rate *Physical Realism* before interpreting the video clips as if they watched the movements first, it would bias the rating. However, the fact that participants rated pictures of the actor and of the animated character before interpreting the material may have affected the results. It can be argued that participants were made aware that it was a comparative study before watching the material and that this awareness could have biased their answers. However the experiment was advertised as a study on the recognition of emotional body language and the aim only explained during debrief. Moreover, several participants thought that they had been asked to interpret the exact same video clips more than once. This suggests that they were not aware of the comparative hypothesis during the test.

After rating the static pictures on screen, participants started watching the presented emotion clips which were played only once and in a randomised order. The participants rated it before clicking to see the next clip. Once all presented emotion clips were interpreted, the post-study questionnaire automatically started. Even though the participants could complete the study at their own pace, the whole procedure took about one hour.

## 3.5. Results

### 3.5.1. Data Confirmation

Before conducting the analysis the Research Corpus was examined to confirm that it adhered to the expected properties.

| | Modality | Presentation Condition | |
|---|---|---|---|
| | | Acted Videos | Animated Videos |
| **Believability** | **Voice** Max=7 | M=4.77 SD=1.09 | M=4.2 SD=0.98 |
| | **No Voice** Max=7 | M=5.33 SD=0.81 | M=4.70 M=1.03 |
| **Naturalness** | **Voice** Max=7 | M=5.00 SD=1.07 | M=4.34 SD=0.87 |
| | **No Voice** Max=7 | M=5.64 SD=0.87 | M=4.83 SD=1.05 |
| **Perceived Strength** | **Voice** Max=5 | M=3.58 SD=0.67 | M=3.24 SD=0.70 |
| | **No Voice** Max=5 | M=3.76 SD=0.53 | M=3.50 SD=0.63 |
| **Number of Correct Interpretations** | **Voice** max=20 | M=8.54 SD=1.94 | M=8.08 SD=2.1 |
| | **No Voice** max=20 | M=5.83 SD=2.85 | M=5.50 SD=2.15 |
| **Number of Correct Interpretations (Negative)** | **Voice** max=10 | M=6.08 SD=1.32 | M=5.08 SD=1.66 |
| | **No Voice** max=10 | M=3.25 SD=1.215 | M=2.17 SD=1.115 |
| **Number of Correct Interpretations (Positive)** | **Voice** max=10 | M=2.46 SD=1.39 | M=3 SD=1 |
| | **No Voice** max=10 | M=2.58 SD=1.78 | M=3.33 SD=1.67 |

**Table 3.3.** Means and Standard deviations of the results.

The recognition rates were calculated to confirm that overall, participants were better than chance level to identify the emotions displayed (table 3.3).

Participants' mean score at the *Emotional Intelligence Questionnaire* was 111.92 out of 150 (SD=10.65). An average of 2.4 out of 7 (SD=1.47) was reported on *Experience in Video Games* and 3.8 out of 7 (SD=1.82) on *Familiarity with Animated Characters*. The mean rate of *Physical Realism* was 5.52 out of 7 (SD=1.83) for the actor and 2.76 out of 7 (SD=1.56) for the animated character (Maximum being 7).

Prior to any analysis, one-way ANOVAs were computed to check if differences confounded the results between the two groups (Voice Vs. No Voice). There was no difference between groups in *Emotional Intelligence* ($F(1,23)=0.61$, $p=0.44$), *Experience in Games* ($F(1,18)=1.53$, $p=0.23$) nor in *Familiarity with Animated Characters* ($F(1,18)=1.53$, $p=0.47$).

### 3.5.2. Physical Realism

Repeated-measures ANOVA were carried out on the data. Assumptions of normality, homogeneity of variance and sphericity were met. Results showed that differences between conditions were unlikely to have arisen by sampling error.

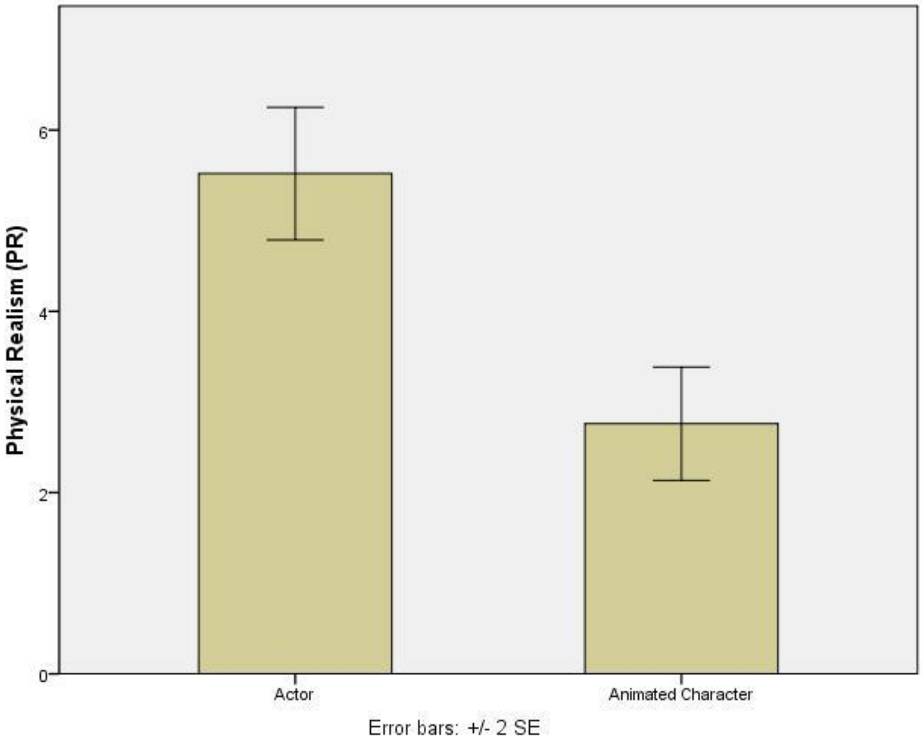


**Figure 3.5.** Physical Realism of the Actor and of the Animated Character.

There was a significant difference in *Physical Realism* ($F(1,23)=49.37$, $p<0.01$) of the static images of the actor and of the animated character. The Actor was perceived as significantly more realistic than the Animated Character (Figure 3.5)

### 3.5.3. Number of correct interpretations

| Emotion Displayed | Modality | Number of Correct Interpretations in the Acted Condition | Number of Correct Interpretations in the animated condition |
|---|---|---|---|
| **Anger** | Voice | Natural: 13/13<br>Stylised: 10/13 | Natural: 12/13<br>Stylised: 9/13 |
| | No Voice | Natural: 3/12<br>Stylised: 10/12 | Natural: 1/12<br>Stylised: 9/12 |
| **Disgust** | Voice | Natural: 12/13<br>Stylised: 11/13 | Natural: 13/13<br>Stylised: 9/13 |
| | No Voice | Natural: 0<br>Stylised: 2/12 | Natural: 2/12<br>Stylised: 1/12 |
| **Shame** | Voice | Natural: 0<br>Stylised: 1/13 | Natural: 1/13<br>Stylised: 3/13 |
| | No Voice | Natural: 0<br>Stylised: 3/12 | Natural: 0<br>Stylised: 0 |
| **Fear** | Voice | Natural: 7/13<br>Stylised: 3/13 | Natural: 6/13<br>Stylised: 2/13 |
| | No Voice | Natural: 7/12<br>Stylised: 3/12 | Natural: 4/12<br>Stylised: 1/12 |
| **Sadness** | Voice | Natural: 11/13<br>Stylised: 9/13 | Natural: 4/13<br>Stylised: 7/13 |
| | No Voice | Natural: 4/12<br>Stylised: 7/12 | Natural: 3/12<br>Stylised: 5/12 |
| **Surprise** | Voice | Natural: 4/13<br>Stylised: 4/13 | Natural: 3/13<br>Stylised: 0 |
| | No Voice | Natural: 3/12<br>Stylised: 2/12 | Natural: 2/12<br>Stylised: 1/12 |
| **Relief** | Voice | Natural: 5/13<br>Stylised: 5/13 | Natural: 6/13<br>Stylised: 9/13 |
| | No Voice | Natural: 2/12<br>Stylised: 5/12 | Natural: 5/12<br>Stylised: 6/12 |
| **Happiness** | Voice | Natural: 1/13<br>Stylised: 2/13 | Natural: 4/13<br>Stylised: 2/13 |
| | No Voice | Natural: 1/12<br>Stylised: 4/12 | Natural: 0/12<br>Stylised: 7/12 |
| **Pride** | Voice | Natural: 0/13<br>Stylised: 3/13 | Natural: 0/13<br>Stylised: 7/13 |
| | No Voice | Natural: 0/12<br>Stylised: 4/12 | Natural: 1/12<br>Stylised: 6/12 |
| **Excitement** | Voice | Natural: 2/13<br>Stylised: 6/13 | Natural: 0/13<br>Stylised: 8/13 |
| | No Voice | Natural: 4/12<br>Stylised: 6/12 | Natural: 5/12<br>Stylised: 7/12 |

**Table 3.4.** Proportion of *Correct Interpretations* per emotions and conditions (13 participants in the Voice condition and 12 in the no voice condition). Chance level would be 1/22.

As shown in Table 3.4, recognition rates for each presented emotion emotion clip, depended on emotion and display conditions (See Confusion Matrices; Table 3.5-3.8), although most of the presented emotion clips were recognised above chance level.

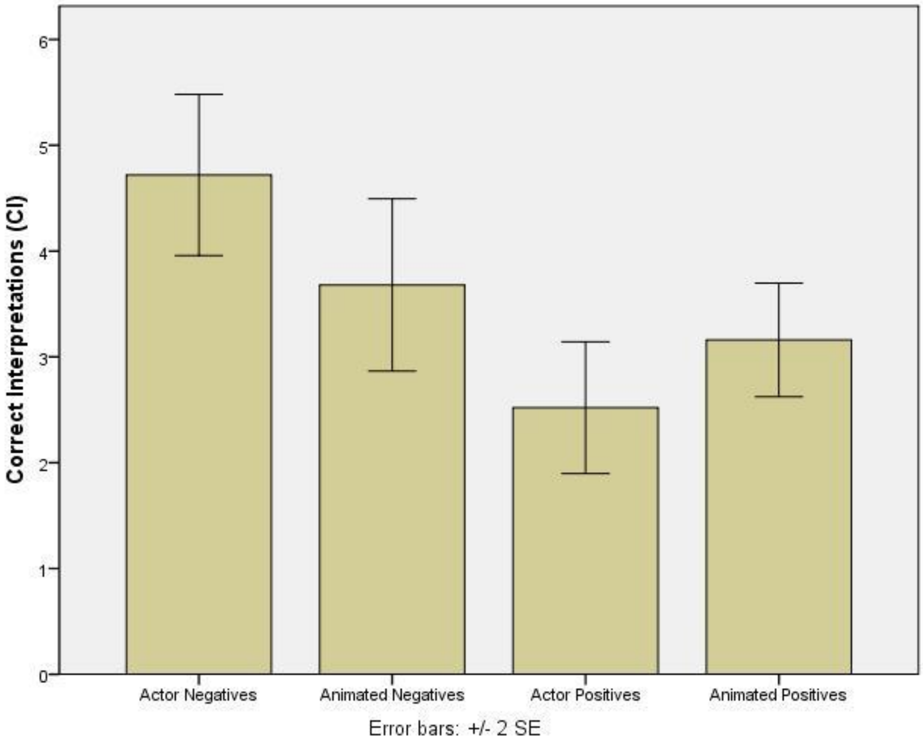


**Figure 3.6.** Means for *Correct Interpretation*

Repeated measures ANOVA did not show a significant difference in the *Number of Correct Interpretations* (F(1,23)=0.9, p=0.4) for the acted and the animated condition, but there was a significant between subjects' effect for *Modality* (F(1,23)=10.7, p<0.01).

However, grouping the displays by valence revealed a significant difference for the negative ones (F(1,23)=14.5, p<0.01) for which participants were more accurate in interpreting the actor (Figure 3.6). There was no such difference for the positive displays (F(1,23)=0.9, p=0.4).

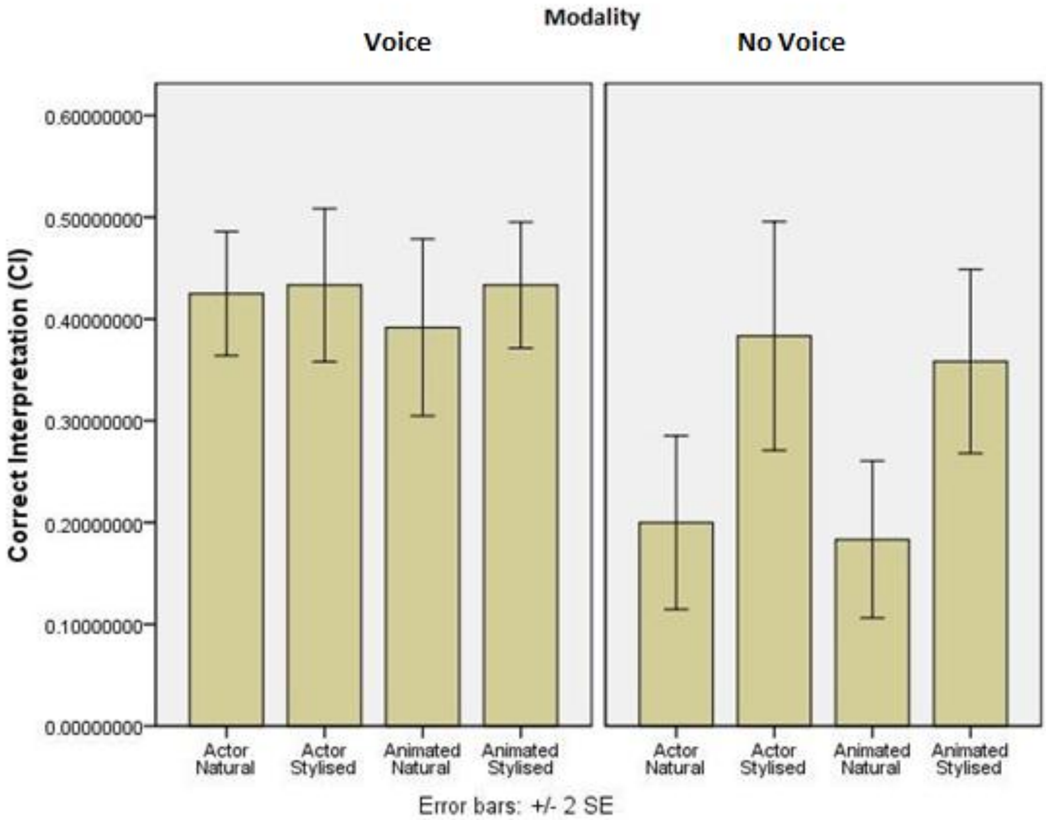


**Figure 3.7.** Means for *Correct Interpretations* for each *Character Type, Action Style* and *Modality*.

Moreover, there was a significant effect of *Modality* on the number of *Correct Interpretations* ($F(1,22)=11.06$, $p<0.01$). The results showed no interaction between *Modality* and *Character Type* ($F(1,22)=0.01$, $p=0.92$), it can therefore be concluded that emotions displayed with *voice* were interpreted more accurately than when displayed without (Figure 3.7).

There was a significant effect of *Action Style* on the number of *Correct Interpretations* ($F(1,22)=13.94$, $p<0.01$). The results also indicate an interaction between *Action Style* and *Modality*. The interaction show that the difference that was found between the Stylised and Natural displays of emotions was present in the No sound condition only (Figure 3.7) and that it was difficult for participants to correctly interpret the natural displays without vocal support.

These results indicate that overall the presented emotional clips' recognition rates were similar when displayed by the Actor or by the Animated Character. They were also better recognised when displayed with voice. However, participants were better in interpreting the negative emotions when displayed by the Actor than when displayed by the Animated Character (Figure 3.7).

The differences in interpretations were further investigated using confusion matrices for the two *Characters Type* and *Action Style*. Since the aim is to compare the interpretation of the emotion displayed by the Animated Character and the Actor, it was decided to group the *Modality* condition.

| **PARTICIPANTS INTERPRETATION** | **EMOTION DISPLAYED** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Excitement** | **Pride** | **Happiness** | **Relief** | **Surprise** | **Sadness** | **Fear** | **Shame** | **Disgust** | **Anger** |
| **Interest** | | 4% | 8% | | 8% | | 4% | | | 4% |
| **Excitement** | 24% | | 8% | | 16% | | | | | |
| **Pride** | | | 4% | | | | | | | |
| **Happiness** | | | 8% | | | | | | | |
| **Enjoyment** | 4% | | 8% | | | | | | | |
| **Tenderness** | | | | | | | 4% | | | |
| **Wonderment** | 12% | | | 4% | 8% | | | | | |
| **Relief** | | 4% | 20% | 28% | 4% | | | 4% | | |
| **Surprise** | 8% | | | | 28% | | 12% | | | |
| **Nostalgia** | | | | | | | | | | |
| **Pity** | | | | 8% | | | | 8% | 4% | |
| **Sadness** | | 20% | 4% | | | 60% | | 36% | 4% | |
| **Fear** | | | | | | | 56% | | | |
| **Shame** | | | 4% | | | 12% | 8% | | | 4% |
| **Guilt** | | | | | | 8% | 4% | 4% | | |
| **Disappointment** | 4% | 52% | 8% | 48% | | 16% | 4% | 24% | 20% | |
| **Jealousy** | 8% | | | | | | | | 4% | |
| **Disgust** | 12% | | | | 4% | | | 4% | 48% | 4% |
| **Contempt** | | 4% | 12% | 8% | 8% | | | 8% | 16% | 12% |
| **Anger** | 28% | 4% | | | 8% | | | | | 64% |
| **Different Emotion** | | 4% | 12% | 4% | 12% | | 8% | 8% | 4% | |
| **No Emotion** | | 4% | 4% | | 4% | 4% | | 4% | | 12% |

**Table 3.5.** Confusion Matrix of the Actor Natural displays

The confusion matrix of the actor natural displays suggest that the positive displays were poorly recognised in this condition (Table 3.5). Moreover, all the positive emotions were often interpreted as negative (Table 3.5). For instance, pride ordinary was interpreted as disappointment by 52% of the participants and as sadness by another 20%.

| PARTICIPANTS INTERPRETATION | EMOTION DISPLAYED | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Excitement** | **Pride** | **Happiness** | **Relief** | **Surprise** | **Sadness** | **Fear** | **Shame** | **Disgust** | **Anger** |
| **Interest** | | | 8% | | 16% | | 8% | | | 8% |
| **Excitement** | 20% | | 8% | | 12% | | | | | |
| **Pride** | | 4% | 8% | | 4% | 4% | | | | 16% |
| **Happiness** | | | 16% | 4% | | | | | | |
| **Enjoyment** | | | 8% | 4% | 4% | | 4% | | | |
| **Tenderness** | | | | | | | 4% | | 4% | |
| **Wonderment** | 36% | | 8% | 4% | 16% | 4% | 4% | | | |
| **Relief** | | 12% | 8% | 44% | 4% | | | | | |
| **Surprise** | 12% | | 4% | | 20% | | 8% | 4% | | |
| **Nostalgia** | | | | | | 4% | | | | |
| **Pity** | | | | | | | | 4% | | |
| **Sadness** | | 4% | | 8% | | 28% | | | | |
| **Fear** | 4% | | 4% | | | | 40% | | | 4% |
| **Shame** | | | | | | 8% | 8% | 4% | 4% | |
| **Guilt** | | | | 16% | | 4% | 4% | 4% | | |
| **Disappointment** | 8% | 24% | | 16% | | 16% | 4% | 24% | 8% | |
| **Jealousy** | 8% | | 4% | | | | | | | |
| **Disgust** | 8% | 4% | 8% | | 4% | | 4% | 4% | 60% | 4% |
| **Contempt** | | 4% | 4% | | 4% | | | 4% | 20% | 4% |
| **Anger** | 4% | 8% | 8% | | 12% | 4% | | | | 52% |
| | | | | | | | | | | |
| **Different Emotion** | | 8% | 4% | | 4% | 8% | 8% | 20% | | 8% |
| **No Emotion** | | 32% | 4% | 20% | | 20% | 4% | 32% | 4% | 4% |

**Table 3.6.** Confusion Matrix of the Animated Natural displays

Overall, the confustion matrix of the animated natural displays suggest a similar pattern as the acted displays. Most of the positive emotions were often interpreted as negative emotions (Table 3.6). However, their exact interpretations were different. For instance, pride was interpreted as no emotion by 32% of the participants and as disappointment by 24% of the participants.

| PARTICIPANTS INTERPRETATION | EMOTION DISPLAYED | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Excitement** | **Pride** | **Happiness** | **Relief** | **Surprise** | **Sadness** | **Fear** | **Shame** | **Disgust** | **Anger** |
| **Interest** | | 4% | | | | | | 4% | 8% | |
| **Excitement** | 48% | 16% | 28% | | 16% | | | | 4% | |
| **Pride** | 4% | 28% | 12% | | | | | | | |
| **Happiness** | 40% | 16% | 24% | | 4% | | | | | |
| **Enjoyment** | 8% | 24% | 24% | | 4% | | | | | |
| **Tenderness** | | | | | | | | | | |
| **Wonderment** | | 4% | 8% | 8% | 12% | | 4% | | 4% | |
| **Relief** | | | | 40% | 12% | | | | | |
| **Surprise** | | | | 12% | 24% | | 8% | | 4% | |
| **Nostalgia** | | | | 8% | | | | | | |
| **Pity** | | | | | | | | | | |
| **Sadness** | | | | | | 64% | | 12% | | |
| **Fear** | | | | | | | 24% | | 4% | |
| **Shame** | | | | | 8% | 12% | 8% | 28% | | |
| **Guilt** | | | | | | 4% | 8% | 32% | | |
| **Disappointment** | | | | 8% | 4% | 20% | 4% | 12% | | |
| **Jealousy** | | 8% | | | | | | | 4% | |
| **Disgust** | | | | 4% | 4% | | 8% | 4% | 52% | |
| **Contempt** | | | | 4% | 8% | | | | | 12% |
| **Anger** | | | | | 4% | | 12% | | 8% | 80% |
| **Different Emotion** | | | 4% | 12% | | | 12% | 8% | 4% | 8% |
| **No Emotion** | | | | 4% | | | 12% | | 8% | |

**Table 3.7.** Confusion Matrix of the Actor Stylised displays

The confusion matrix of the actor stylised displays highlight that overall stylised positive emotions were interpreted as positive and, similarly, negative displays were interpreted as negative (Table 3.7). For instance happiness and excitement were often misinterpreted for one another.

| PARTICIPANTS INTERPRETATION | EMOTION DISPLAYED | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Excitement** | **Pride** | **Happiness** | **Relief** | **Surprise** | **Sadness** | **Fear** | **Shame** | **Disgust** | **Anger** |
| **Interest** | | 4% | 4% | | 4% | | | 12% | 16% | |
| **Excitement** | 60% | 12% | 44% | | 24% | | 8% | 4% | | |
| **Pride** | 8% | 52% | 4% | | | | | | | |
| **Happiness** | 16% | 4% | 36% | | 8% | | | 8% | | |
| **Enjoyment** | 12% | 8% | 8% | | 4% | | | | | |
| **Tenderness** | | | | | | | | | 4% | |
| **Wonderment** | 4% | | 4% | | 12% | | 8% | | | |
| **Relief** | | | | 60% | | | | 4% | | |
| **Surprise** | | | | 24% | 4% | | 16% | 4% | | |
| **Nostalgia** | | 4% | | | | | | | | |
| **Pity** | | | | | | 4% | | | | |
| **Sadness** | | | | | | 48% | | 20% | 4% | |
| **Fear** | | | | | | | 12% | | | |
| **Shame** | | | | | | 12% | | 12% | | |
| **Guilt** | | | | | | | 8% | 20% | | |
| **Disappointment** | | | | 8% | 4% | 36% | | 12% | 4% | 4% |
| **Jealousy** | | 4% | | | 4% | | | | 4% | |
| **Disgust** | | | | | | | 12% | | 40% | |
| **Contempt** | | | | | 12% | | 8% | | 4% | 12% |
| **Anger** | | 4% | | | 20% | | 16% | | 4% | 72% |
| **Different Emotion** | | 8% | | 8% | 4% | | 12% | 4% | 12% | |
| **No Emotion** | | | | | | | | | 8% | 12% |

**Table 3.8.** Confusion Matrix of the Animated Stylised displays

The confusion matrix of the animated stylised displays (Table 3.8) show similar patterns of interpretation as the acted ones (Table 3.7). Overall stylised positive emotions were interpreted as positive and, similarly, negative displays were interpreted as negative. For instance happiness and excitement were often inversed in this condition as well.

### 3.5.4. Emotional Strength

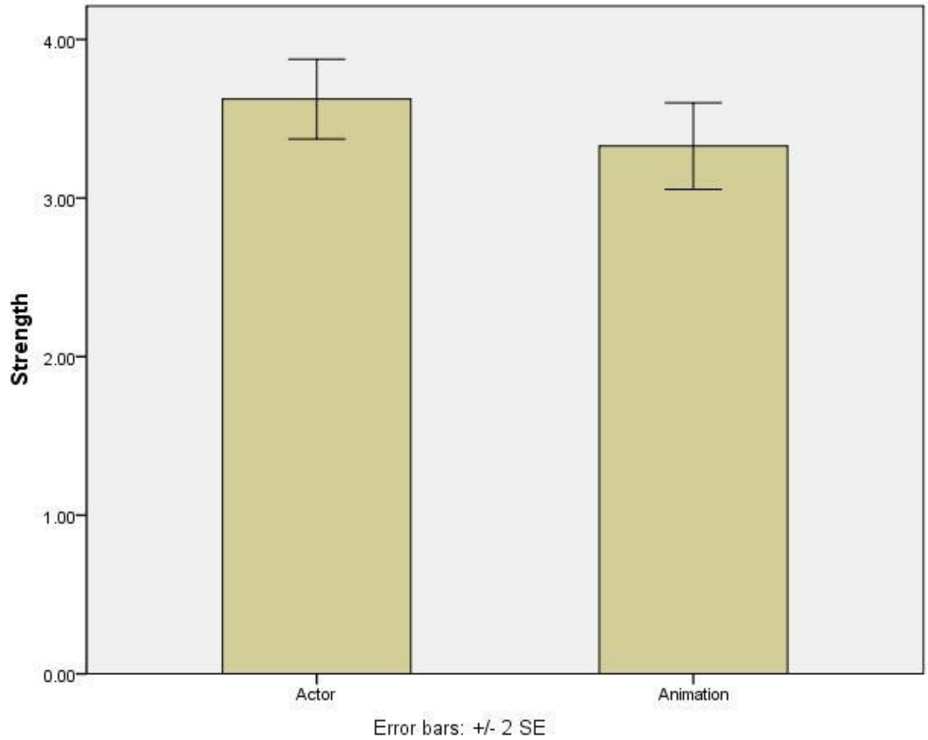


**Figure 3.8.** Emotional *Strength* for each *Character Type*,

There was a significant difference in *Strength* between the actor and the Animated Character ($F(1,23)=42.87$, $p<0.01$). Overall emotions were perceived as Stronger (*Strength*) when displayed by the Actor than when displayed by the Animated Character (Figure 3.8). However, *Modality* had no significant effect on *Strength* ($F(1,22)=0.74$, $p=0.40$).

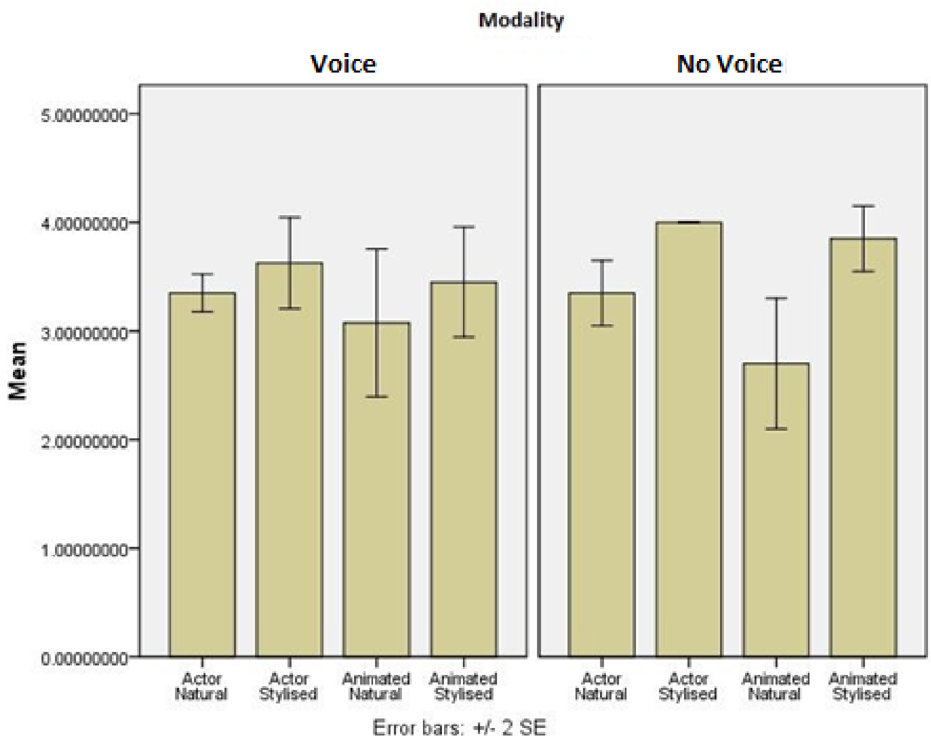


**Figure 3.9.** Emotional *Strength* for each *Character Type*, *Action Style* and *Modality*.

*Action Style* had a significant effect on *Strength* ($F(1,23)=14.10$, $p<0.05$). Emotions were perceived as stronger when displayed in the Stylised version (Figure 3.9). There was no interaction between *Action Style* and *Character Type* ($F(1,4)=0.12$, $p=0.85$). This suggests that the effect of *Action Style* was similar for the animated character and for the actor.

### 3.5.5. Believability

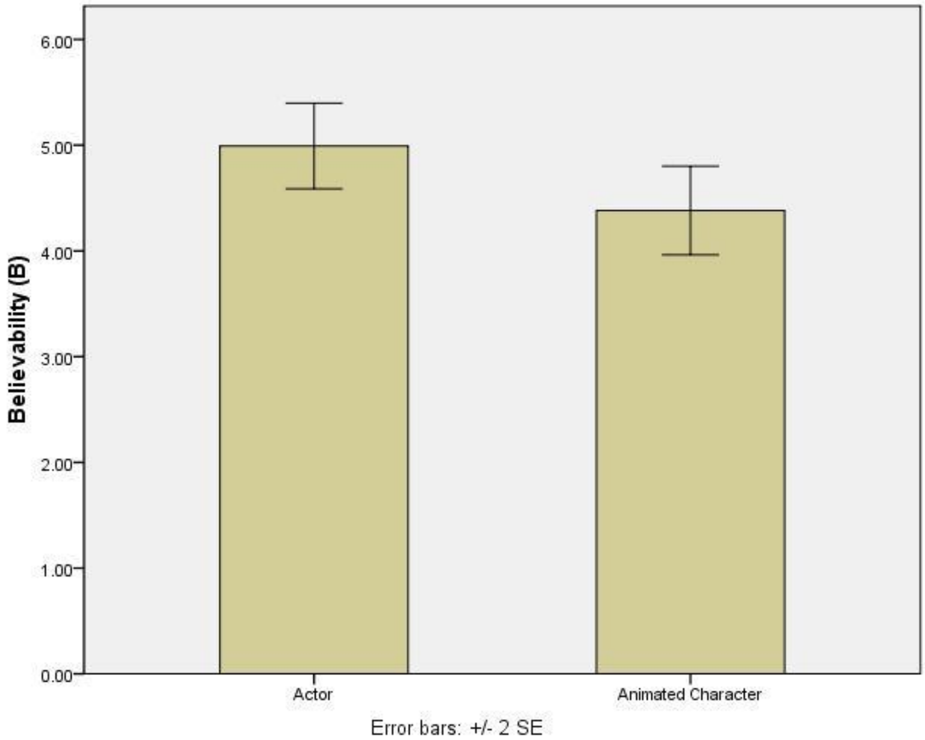


**Figure 3.10.** *Believability* of the Actor and of the Animated Character

There was a significant difference in *Believability* between the Actor and the Animated Character (F(1,23)=11.24, p<0.01). Emotions were found more *Believable* when displayed by the Actor than when displayed by the Animated Character (Figure 3.10). However, *Modality* had no significant effect on *Believability* (F(1,22)=2.18, p=0.15).

Moreover, there was no effect of *Action Style* on *Believability* (F(1,23)=0.37, p=0.55).

### 3.5.6. Naturalness

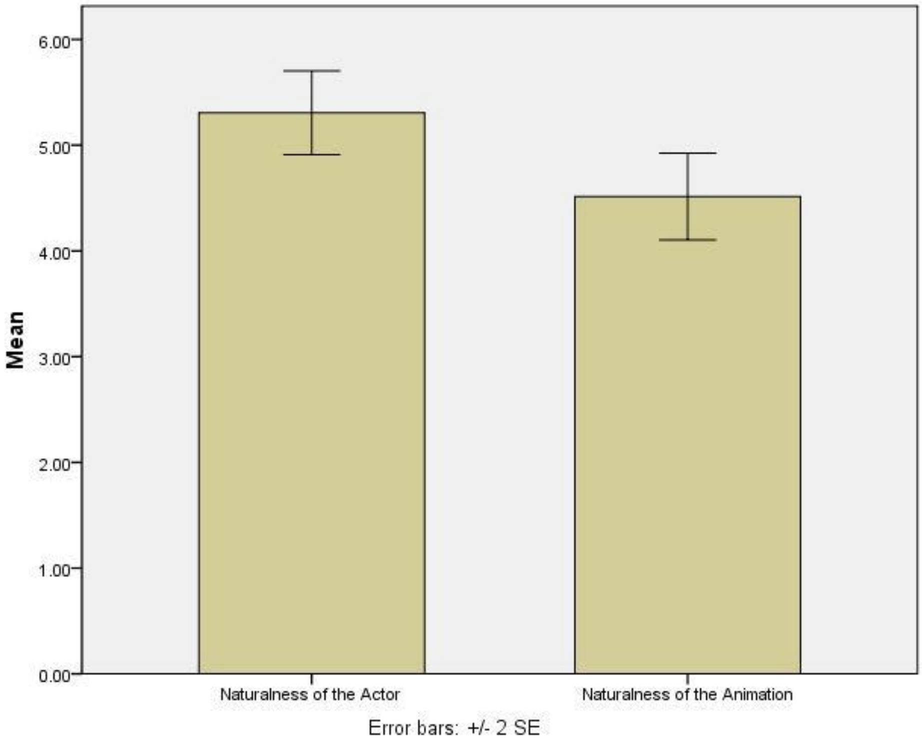


**Figure 3.11.** *Naturalness* of the Actor and of the Animated Character

There was a significant difference in *Naturalness* between the Actor and the Animated Character (F(1,23)=23.51, p<0.01). Emotions were found more *Natural* when displayed by the Actor than when displayed by the Animated Character (Figure 3.11). However, *Modality* had no significant effect on *Naturalness* (F(1,22)=2.37, p=0.14).

Moreover, there was no effect of *Action Style* on *Naturalness* (F(1,23)=0.12, p=0.74).

### 3.5.7. Individual Differences

To assess the correlation of *Emotional Intelligence* with the *Number of Correct Interpretations*, Pearson correlations were computed. There was no correlation between *Emotional Intelligence* and the *Number of Correct Interpretations* for the actor (r(23)=-0.3, p=0.09) nor for the animated character (r(23)=-0.2, p=0.1). This indicates that *Emotional Intelligence*, as measured by the questionnaire used does not co-vary with participants skills to identify the emotional body language displayed in this study.

Similarly Pearson correlations were computed for *Emotional Intelligence* and *Believability*. There was no relation between *Emotional Intelligence* and the *Believability* of the actor ($r(23)=0.27$, $p=0.10$). However, *Emotional Intelligence* was positively correlated with the *Believability* of the animated character ($r(23)=0.48$, $p<0.01$). The results are similar with the perceived *Naturalness*, as there was no correlation between *Emotional Intelligence* and the *Naturalness* ($r(23)=0.2$, $p=0.3$) of the actor. However, *Emotional Intelligence* was again positively correlated with the *Naturalness* of the animated character ($r(23)=0.5$, $p<0.01$). This indicates that participants' with a high score in *Emotional Intelligence* tended to find the emotions displayed by the animated character more believable and more natural than participants who scored lower.

Pearson's test also showed a positive correlation between *Experience in Video Games* and the *Number of Correct Interpretations* for the actor ($r(18)=0.37$, $p<0.05$) but not for the animated character ($r(18)=0.17$, $p=0.24$). However, *Familiarity with Animated Characters* was positively correlated with the *Number of Correct Interpretations* for the actor ($r(18)=0.64$, $p<0.01$) as well as for the animation ($r(18)=0.38$, $p=0.05$).

There was a positive correlation between *Physical Realism of the Actor* with *Believability* of the actor ($r(23)=0.49$, $p<0.01$) and with *Naturalness* of the actor ($r(23)=0.60$, $p<0.01$). However, there was no such correlation between *Physical Realism of the Animated Character* and *Believability* of the animated character ($r(23)=0.04$, $p=0.42$) nor with *naturalness* of the animated character ($r(23)=0.245$, $p=0.12$).

There was a positive correlation between *Experience in Games* and *Physical Realism of the Actor* ($r(23)=0.39$, $p<0.05$) as well as with *Physical Realism of the Animated Character* ($r(23)=0.75$, $p<0.01$). Similarly *Familiarity with Animated Characters* was positively correlated to *Physical Realism of the Actor* ($r(23)=045$, $p<0.05$) and to *Physical Realism of the Animated Character* ($r(23)=0.54$, $p<0.01$).

Finally, four participants out of twenty-five reported not noticing the presence of animations. Only one participant reported being disturbed by the presence of animation.

# 3.6. Discussions

## 3.6.1. Interpretations of the Displays

The results showed that overall, there was no difference in the number of displays that were correctly interpreted between the actor and the animated character (Table 3.3). Moreover, *Modality* had a positive effect on the number of *Correct Interpretations*. Emotions displayed with the context of the voice were easier for participants to identify. This is certainly due to the additional information provided by the voice but also to the emotional information conveyed by the speech itself (i.e. intonation, prosody).

Although it could be due to a lack of statistical power, the lack of difference in *Correct Interpretation* between the actor and the animated character (Section 3.5.3) is a very encouraging result for the use of Virtual Environments for social task training. If it can be verified, it would suggest that emotions displayed by a human or an animated character would be interpreted similarly within a Virtual Environment and therefore could be used for training.

However, it cannot be concluded that the interpretation of an actor and an animated character is completely similar, as grouping the displays by valence shows that for the negative emotions, participants were better at interpreting the actor, whilst there was no such difference for the positive emotions. It is important to emphasize, that these differences occurred between videos of an actor performing and the animated movements that were captured from the same performance.

The fact that no difference was found in the number of correct interpretations (Section 3.5.3) is encouraging, however, it may also be due to a floor effect (i.e. the results being too low to find any difference), on the positive displays (i.e. Excitement, Happiness, Pride, Positive Surprise, Relief). The recognition rate of these videos was very low in both conditions (Table 3.4). Therefore, this result may also be due to the fact that it was impossible for participants to distinguish between so many positive emotional states just from emotional body language, even with voice as well. However, this low recognition rate is consistent with Ekman's basic theory of emotion (Ekman, 1992), which considers that happiness is at the core of all positive emotions. The basic theory of emotion thus would suggest that the positive displays will be very hard to accurately distinguish because of their similarity.

It should be noted that the floor effect could also be due to poor performances by the actor for the positive displays. Although, the recognition rate is above chance level, the confusion matrixes suggest that, at least for the ordinary displays, the positive emotions are often confused with negative ones (Table 3.5 and Table 3.6). Such a problem with performances was not anticipated as existing studies suggest that it is possible to interpreted positive emotions watching emotional body language only (Section 1.4). There are many possibilities for this, including the actor being aware of the experimental conditions and changing his acting to match the condition, resulting in ordinary displays being confounded.

### 3.6.2. Effect of Character Type on Strength, Believability and Naturalness

The animated character has been perceived as less natural (*Naturalness*), less believable (*Believability*) and less emotional (*Strength*). This could be due to the Motion Capture technology used as it fails to capture secondary animations such as finger movements and micro-gestures, such as skin deformation. This could also explain why participants' interpreted more accurately the actor displaying a negative emotion than the animated character. This is suggested by the significant difference between the actor and the animated character for the presented emotion clips where a negative emotion (i.e. Anger, Fear, Disgust, Sadness, Shame) was expressed. Thus it could be suggested that these differences are due to secondary gestures that were not captured by the Motion Capture technology. If this is the case, Virtual Environments for social task training using animated characters would have to procedurally or manually reintroduce these missing secondary animations in order for the experience to be comparable with real-life ones.

However, these differences could also be due to the position of the Animated Character on the Uncanny Valley. The significant difference that was found in *Physical Realism* between the actor and the animated character indicates that the animated character is situated on the left of the actor (Figure 3.1). Therefore, the differences in *Strength*, *Believability* and *Naturalness* could also indicate that the animated character 'fell' in the Uncanny Valley. If these differences are due to the Uncanny Valley, it would imply that such characters are not suitable for social task training. A Virtual Environment for social task training needs to induce presence and this would not be possible if the animated characters are systematically perceived as unbelievable. Moreover, the difference in interpretation could also be counterproductive for social task training as it would certainly result in learning skills and simulated situations not comparable to real life ones.

It is therefore an important issue that will have to be fully investigated (Chapter 4) as it is currently not possible to establish whether these differences are due to the gestures that were not recreated by the Motion Capture Technology or to an Uncanny Valley effect as they could both result in these perceptual differences.

Interestingly, *Modality* had no effect on *Believability* or on *Naturalness*. This could indicate that the voice of the actor was appropriate for both the actor and the animated character. Nevertheless, it could also be due to the low number of participants in the experimental between participants condition (Table 3.2).

### 3.6.3. Individual Differences are related to the perception of the displays

Interestingly, participants who scored high at the *Emotional Intelligence* questionnaire were less affected by the drop in *Believability* and *Naturalness* of the animated displays. This was shown by the positive correlation between *Emotional Intelligence*, *Believability* and *Naturalness* of the animation (Section 3.5.7). However, the lack of a correlation between *Emotional Intelligence* and the number of correct interpretations could indicate that either the questionnaire failed to assess participants' ability to accurately interpret emotional body language or that the videos presented lacked the appropriate emotional cues. The second possibility can be dismissed by the fact that for almost all emotional clips the recognition rate is above chance level (Table 3.2).

The positive correlation between familiarity with animated characters and the number of correct interpretations in both conditions is very encouraging for social task training. It shows that participants who reported being familiar with animated characters were more accurate in interpreting the video clips including the acted ones. However, it is not possible to conclude that these participants learned from previous experience how to interpret emotional body language. The positive correlations between experience in video games and familiarity with animated character with the physical realism of the actor and animated character could simply imply that participants' familiar with animated character were less affected by the appearance of the actor and the animated character, which was different from every day displays: they wore the motion capture suit and their faces were pixelated. This could also explain why participants familiar with animation were more accurate. Therefore, it will be necessary to determine whether it is linked to the appearance. It would also be of interest to investigate whether it is related to specific types of experience with animated characters (different type of animations, interaction…) as this may affect the effectiveness of a Virtual Environment for training.

## 3.7. Conclusion

The results established that it is possible to correctly identify emotional body language animated with motion capture. However, significant differences were found when compared with the video of the actor (see Section 3.6). Nevertheless, these differences were not unexpected and it can be argued that animated characters cannot be, in terms of believability and naturalness, as good as videos of a real person. It can be argued that the animated character used for this study, given current technology, is as realistic as the one that would be used for social task training (Figure 1.4 Top), such as medical interview. These characters have to be animated in real time and thus would also be susceptible to the same drop in believability and naturalness. However, in order to start exploring the shape of the uncanny valley, it will be necessary to use more than two characters. The model predicts differences between characters on both 'sides' of the valley (Figure 3.1), hence at least three characters will be necessary to test its shape. This is further investigated in the next chapters.

It was not possible to clearly establish if the difference between the animated character and the video of the actor was due to the micro-gestures that Motion Capture failed to recreate (Behavioural Realism) or if it was due to the physical appearance of the animated character (Physical realism) and its position on the Uncanny Valley. Indeed, the differences that were found in this study may be due to the inherent knowledge on the part of the viewer, that the character is not real. In other words, although it is possible to identify emotions displayed by animated characters, the differences that were found in this study could raise serious issues regarding their use for social task training.

Hence, the next research step is to investigate whether these effects were due to the change of physical realism of the character (Figure 3.3) or to the loss of secondary animations resulting in lower behavioural realism. It is expected that increasing the level of realism will affect the difference in believability, naturalness and will explore the real shape of the Uncanny Valley.

# Chapter 4 - A Comparative Study of Emotional Body Language Displayed by Humans and Animated Characters: Study Design

## 4.1. Introduction

### 4.1.1. Secondary Animation

The drop found in believability and recognition rate when comparing the same emotional body language simultaneously recorded via video and motion capture (see results reported in Section 3.5) could be due to the motion capture technology used for animation. Indeed, motion capture deteriorates the body language as it fails to record micro gestures and secondary animations such as breathing and hand movements. Breathing is known to participate to the expression of specific emotions (Dantzer, 2005). Hands and fingers movements have been found to contribute to the expression of certain emotions as well (Wallbott, 1998). These types of clues are not recorded by the motion capture system that was used and this could explain the drop in recognition and in believability that was found in Chapter 3. However, it would not be a problem for an animated character in a Virtual Environment as these clues could be procedurally recreated and hence would not be missing.

### 4.1.2. Visual Realism

However, the Uncanny Valley predicts that these differences could also be due to the nature of the animated character itself (Computer generated vs. biological character) and to the difference in physical realism between the actor and the animated character. The Uncanny Valley suggests that realistic characters would not be believable and therefore would not be efficient for social task training. This is why it is important to fully investigate this possibility.

# 4.2. Method

## 4.2.1. Design

To investigate these two possibilities, an experiment which tested the Uncanny Valley as a model was conducted. Due to its shape, the only way to do so is to test whether a highly realistic character is less believable (i.e. falls into the Uncanny Valley) than a non-realistic character when they both display the same emotional body language. Nevertheless, the shape of the Uncanny Valley limits such study. If an uncanny effect is not found it could always mean that the animate characters tested are situated on the right or left part of the valley rather than the model being incorrect. Moreover, the claim that the Uncanny Valley results from high realism (but not high enough) has still to be tested. Therefore, the experiment was designed to test the Uncanny Valley by investigating the effect of manipulating realism on the interpretations and perception of the display. Realism can be divided in two distinct categories, the level of physical realism (i.e. anatomy) as well as the level of behavioural realism (i.e. the way a character moves). The level of physical realism was manipulated by using different type of characters. The level of behavioural realism was also manipulated by using different action styles. As in Chapter 3, interpretation was defined in terms of emotions identified and in terms of emotional strength of a display. Finally, in order to test the Uncanny Valley, the believability and naturalness of the displays were investigated.

### 4.2.2. Independent Variables

| Character Type / Frame Rate | Actor | Realistic Character | Simplified Character |
|---|---|---|---|
| 12 Frames/Second | Ordinary / Stylised | Ordinary / Stylised | Ordinary / Stylised |
| 25 Frames/Second | Ordinary / Stylised | Ordinary / Stylised | Ordinary / Stylised |

**Table 4.1.** Experiment conditions. The effect of *Character Type* was tested within subjects and *Frame Rate* between subjects.

The experiment used a mixed design. The Independent Variables were defined for the exploration of the Uncanny Valley model. Realism was divided in two distinct categories, the level of physical realism (i.e. anatomy) and the level of behavioural realism (i.e. the way a character moves). Physical realism was manipulated by changing the type of character (*Character Type*) displaying the emotions. The characters varied in their level of similarity to an anatomical accurate form. The most realistic character used was an actor video recorded (Figure 4.1-top), followed by a realistic computer generated character (Figure 4.1-middle) and a simplified one (Figure 4.1-bottom); (Actor Vs. Realistic Character Vs. Simplified Character) (Table 4.1).

The fact that traditional animation usually uses stylised and unrealistic displays of emotions, which are adapted to the physical appearance of the characters (Appendix A), may have an effect at different levels of physical realism. This could affect the perception of the animated display (participants on seeing an animated character could expect stylised and exaggerated display of emotions rather than ordinary displays). Therefore, behavioural realism was manipulated by changing the style of display (*Action Style*). The style of movements varied in their similarity with the ones that could be displayed in the real world and to the ones that can be find in traditional animation (Ordinary Vs. Stylised). The experiment also investigated whether there is an 'uncanny effect' in the display, by comparing the perceived believability and naturalness, of ordinary and stylised displays of emotion *(Action Style)* (Table 4.1).

With respect to the results of the previous study (see Section 3.5), it is unclear whether the movements that motion capture failed to record have an effect on how the emotional body language is perceived. Therefore, an additional independent variable was included in which the quality of the emotional body language displayed was deteriorated. This approach was chosen rather than reintroducing the secondary animations not captured by Motion Capture as it would be impossible to do so without confounding the results. It would be almost impossible to insure that the secondary animations added are exactly the same as in the original performance. However, it is possible to deteriorate the body language displayed in a way that is consistent across all the conditions. This was done by reducing the frame rate from twenty-five frames per second to twelve frames per second by removing every other frame (12FPS Vs 25FPS). Twelve frames per second was chosen as it is considered just enough to maintain a satisfactory illusion of continuous motion (Chapman & Chapman, 2004), therefore creating a condition in which body movements are deteriorated but the illusion of motion maintained (*Frame Rate*).

### 4.2.3. Dependent Variables

The experiment was conducted to investigate whether the problem described by the Uncanny Valley applies to the perception of emotional body language. A prediction based on the Uncanny Valley is that a highly realistic character will be harder to interpret and will also be perceived as less emotional (Chapter 2). To test this prediction, two dependent variables were used to record emotional interpretation. It was defined and recorded in terms of emotional identification (*Correct Identification*) and perceived emotional strength (*Strength*). Moreover, another prediction based on the Uncanny Valley is that as characters get more realistic, they will be subject to a drop in believability and naturalness (see Chapter 2). Believability and naturalness are very important in animation as they contribute to the 'illusion of life' (Thomas & Johnston, 1995). Thus, this prediction was also tested by adding perceived believability (*Believability*) and naturalness (*Naturalness*) as dependent variables.

Following these hypotheses, it was expected that a complex set of interactions between *Character Type*, *Action Style* and *Frame Rate* will be found in the results. More precisely, it was predicted that the effect of *Character Type* will depend on *Action Style* and on *Frame Rate*. For the ordinary displays, it was predicted that the characters' appearance will have an effect and that the actor (consistent realism) will be perceived as better than the simplified character, which in turn will be perceived better than the realistic character (Figure 4.1). Moreover, it was predicted that the actor will be perceived as better when displaying ordinary movements at 25 frames per second whereas the opposite was predicted for the two animations. Similarly, the simplified character should be perceived as better when displaying stylised movements at 12 frames rate per second as it is the frame rate used in classical animation.

In addition to external factors another issue to consider when making emotion judgments is the participants' innate ability to accurately interpret emotional behaviour. Emotional Intelligence was thus recorded to determine whether this correlates with the viewer's ability to classify emotion in either the video or in the animated conditions. Similarly, previous experience with video games and animated characters might affect the way the emotional language is perceived, especially in the animated conditions. It is possible that participants experienced in playing video games or generally very familiar with animation are accustomed to the 'uncanny' effect and will be less affected. Therefore, the study considered potential effects including whether they correlate with participants' perceived believability and naturalness of the emotional body language displayed. No prediction regarding individual differences can be made based on the Uncanny Valley. However, it is expected that they will be related to the way the agents and emotional displays are perceived and interpreted.

### 4.2.4. Research Questions

In summary, four main questions were identified and tested:

(Q1) Does the *Character Type* affect the *Correct Identification*, *Strength*, *Believability* and *Naturalness* of the emotional body language displayed?

(Q2) Does the *Action Style* affect the *Correct Identification*, *Strength*, *Believability* and *Naturalness* of the emotional body language displayed?

(Q3) Does the *Frame Rate* affect the *Correct Identification*, *Strength*, *Believability* and *Naturalness* of the emotional body language displayed?

(Q4) Are personal differences including *Emotional Intelligence*, *Experience in Playing Video Game* or *Familiarity with Animated Character* related to *Correct Identification*, *Strength*, *Believability* and *Naturalness*?

### 4.2.5. Participants

40 Participants were recruited, all members of staff or students of the University of Portsmouth (21 females and 19 males) ranging in age from 20 to 60. Participants were randomly allocated to one of the two groups (12 Frames/Second and 25 Frames/Second in Table 4.1). The twenty participants (12 Females and 8 males) in the 25 Frames/Second Movements condition ranged in age from 21 to 60 (M=35, SD=12). The twenty participants (9 females and 11 males) in the 12 Frames/Second Movements condition ranged in age from 20 to 53 (M=36, SD=11). Participants were entered in a raffle to win an Ipod in exchange for participation.

### 4.2.6. Apparatus

In order to refine and check that the results of the first study were not due to specific gestures within the performances recorded, a different professional actor and a professional director (from study 1 – Chapter 3) were used.

The creation of the displayed clips followed a similar approach as for the previous study (Figure 3.2). However, to avoid a potential confound (Section 3.6.1), and in contrast with the first study, only the director was briefed on the nature of the experiment. This was done to ensure that the actor's performances were not influenced by the Action Style or Character Types conditions, any more than the functional changes required.

Both the actor and the director were trained in "Lecoq" theatrical techniques, which emphasise physicality and strong embodiment of emotion. "Lecoq" was chosen because a large part of the training uses various types of masks. Therefore putting a strong emphasis on bodily expressions of emotions. Prior to the recording, the director prepared a set of small scenarios appropriate for each emotion. Although, in contrast with the first study (Chapter 3) there was no constraint in the type of sentence that could be used as there was no modality (Voice/No Voice) condition.

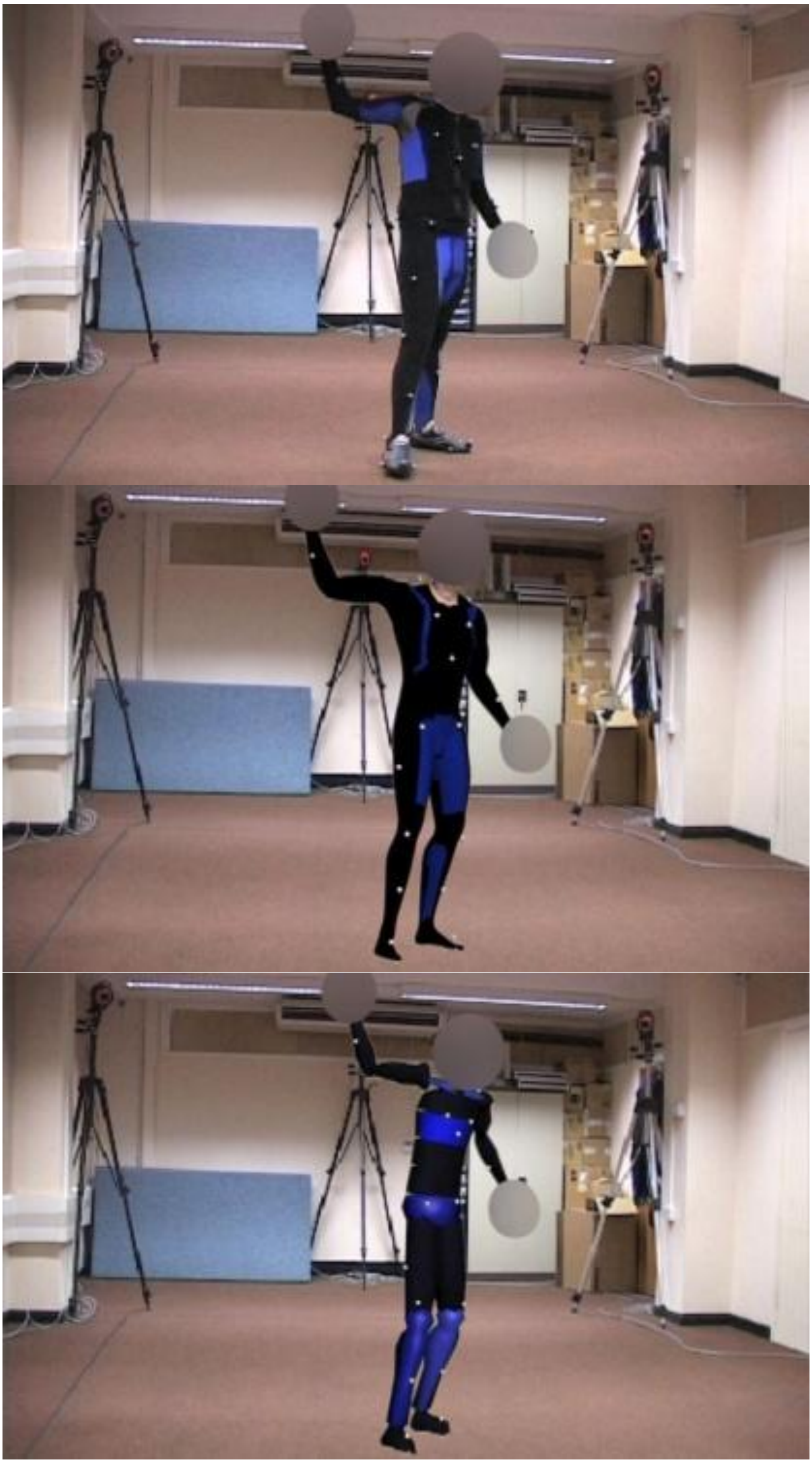

**Figure 4.1.** Screen shots of the Actor, of the Realistic Character (polygon mesh) and of the Simplified Character (anatomic primitives).

The actor performed the same set of ten emotions as in the first study (Chapter 3): Anger, disgust, shame, fear, sadness, positive surprise, relief, happiness, pride and excitement. Each emotion was performed in two different ways, an ordinary version for which the actor was directed so that he would act naturally and an exaggerated or stylised one closer to the style of traditional animation. He also performed a 'neutral' state. During the recording, the director and actor used improvisation techniques to "push" to reach the desired emotion. During the recording, the actor followed the director's instructions. He continued performing each emotion till the director was satisfied with the performance. This ensured that the recorded display conveyed the intended emotion as well as being appropriate for the ordinary or stylised condition.

To ensure equivalency across conditions, the actor was video recorded (Figure 4.1A) and motion captured simultaneously. The videos were recorded using a Sony PD170P, at Standard Definition (525i in 4:3 mode). Motion capture data was recorded using a VICON optical motion capture system (Figure 3.3) (8 "v-series" cameras at VGA, 640x480, resolution, resolved with VICON IQ2.). The motion capture data was visually checked and cleaned when necessary and then used to animate two characters (Figure 4.1) so that they displayed the same body language. Three different characters were used: The actor which was video recorded (Figure 4.4-top), a physically realistic model based on polygon mesh, and a physically simpler character made of simple shapes for each limb component.

Autodesk 3DS Max was used to rig the animated characters and to render the animations. When rendering the animations, a green background, equivalent to a chroma key, was placed behind the animated character. It was then overlaid with a video recording of the room used to record the motion capture using Adobe After Effects. Due to the complexity of the lighting, almost no shadows were present in the actor condition (Figure 4.1). No shadows were rendered in the animated condition. The faces and hands of the actor and of the animated characters were pixelated (using Adobe After Effects as well), so that this source of information was removed along with the possible uncanny effect that may come from facial animations (Figure 4.1). In addition, to remove possible effects, such as differences in dress of the actor and animated character, the three appeared in a motion capture suit (Figure 4.1), were sized to look similar (faces were not visible) and were put in the same visual context (Figure 4.1). Thus, sixty-three presented emotion clips were created: (10 emotions x 2 Style

(Natural and Stylised) + 1 neutral state) x 3 Character Types (Actor, Realistic Character and Simplified Character).

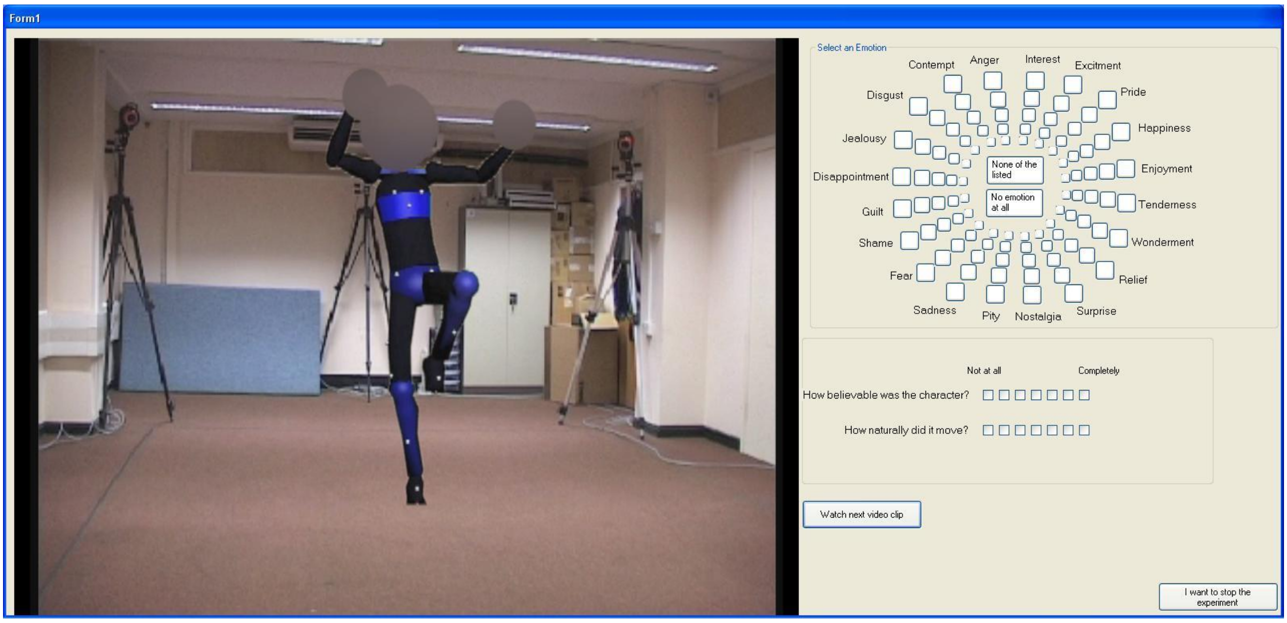


**Figure 4.2.** Screen shots of the application

To record participants' answers, the material was embedded into a custom made interface, which was used for displaying the presented emotion clips at life size as well as recording the participants' answers (Figure 4.2).

The software used to display the material did not record the order in which the clips were displayed. It was therefore not possible to test whether the order in which participants saw the emotional displays had an effect on their interpretation. Any side effect such as tiredness could not be accounted for. However, it should be noted that participants did not report being tired or finding the study too long during the debrief session.

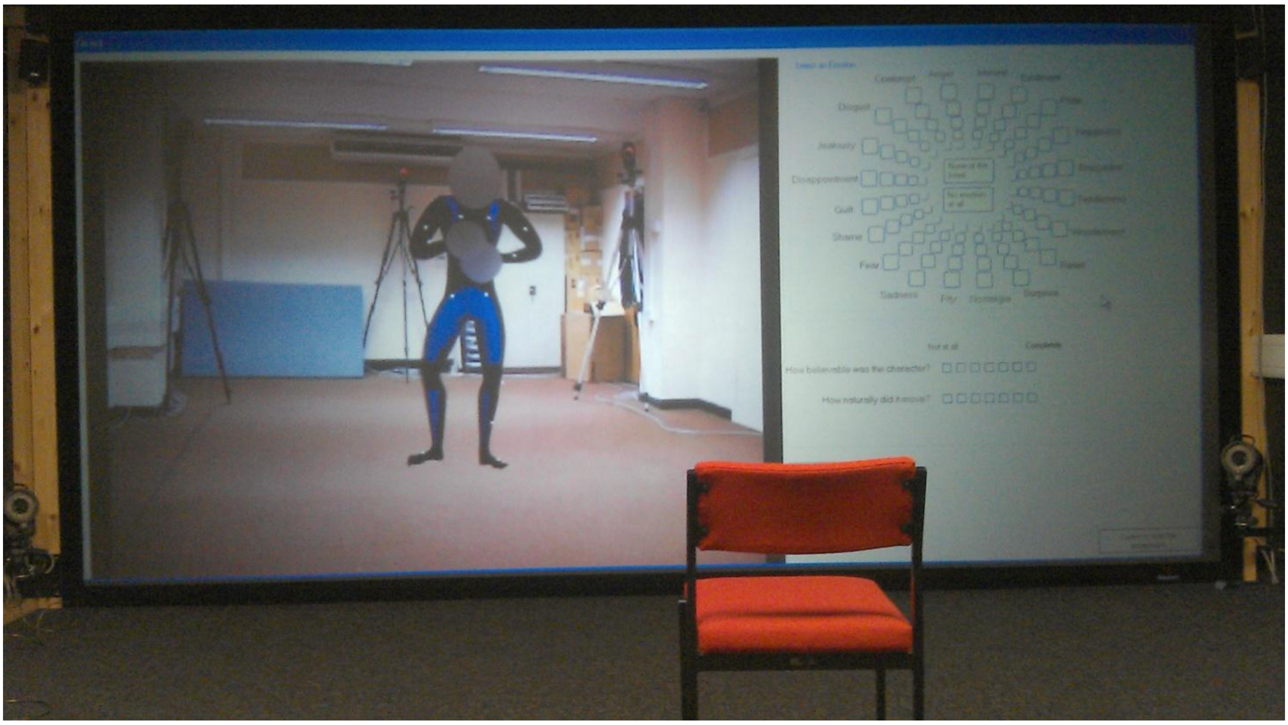

**Figure 4.3.** Photo of the experimental setting. The screen dimensions are 5x2.5 meters

The videos, the animations and the questionnaire were displayed on a 5m x 2.5m rear projection screen (Figure 4.3).

### 4.2.7. Materials

In order to record the participants' interpretations (*Correct Identification* and *Strength*) of the emotional body language displayed, an existing questionnaire has been modified. It is based on the Geneva Emotional Wheel (Scherer, 2005) that places twenty emotional labels on a two dimensional axis; valence and control. The Geneva Emotional Wheel is usually used for self-report (i.e. participants reporting their own emotional state). However, in this study, it has been used to report on participants' interpretations of the clips. The wheel's centre includes two additional options, "no emotion at all" (neutral state) and "none of the above", in case a participant perceives an emotion not on the wheel (Figure 4.2). Each participant was also asked to indicate the *Strength* for every presented emotion clip (five point Lickert scale radiating out from the centre).

For each clip, participants were also asked to rate the *Believability* and *Naturalness* on a seven point lickert scale (1="Not at all", 7="Completely"). A Post-study questionnaire gathered personal information including gender (Male, Female), age, *Experience of Video Gaming* and *Familiarity with Animated Characters* on a five point lickert scale (1=”Not at all”, 5=”Very much”). Finally participants were presented with a picture of the Simplified Character, a picture of the Realistic Character and a picture of the Actor and asked, for each of them, if during the study they thought this character was an actor or an animation.

The *Emotional Intelligence* was measured using the 33-item emotional intelligence scale (Schutte et al., 1998).

### 4.2.8. Procedure

The study was advertised as a study on emotional body language, no mention of the presence of animated characters was made to participants prior to the experiment.

All participants were tested by the same experimenter but in individual sessions. Each session began by obtaining consent, followed by the emotional intelligence questionnaire. After completion of the questionnaire, participants were told how to use the software and given an explanation of the emotional wheel. The term believability was clearly defined as “*to what extent do you think the character is feeling the emotion”* and naturalness was defined as *“the quality of the way the character moves”*. Participants were informed that faces and hands were blurred before they watched and assessed the 63 video clips (10 emotions +1 neutral state) * 3 Character Types). Each video was played through only once. Then participants responded, which triggered the next video. When all video clips were interpreted, the post-study questionnaire automatically started. Finally, participants were fully debriefed regarding the purpose of the study. The whole procedure took less than one hour.

# Chapter 5 - A Comparative Study of Emotional Body Language Displayed by Humans and Animated Characters: Study Results

## 5.1. Data Confirmation

To see if differences could have confounded the results between the two separate *Frame Rate* (12 FPS Vs 25 FPS) groups, One-way ANOVAs were carried out on the participant data between. There was no difference among for *Emotional Intelligence* ($F(1,38)=0.01$, $p=0.92$), *Experience in Video Games* ($F(1,38)=0.001$, $p=1.00$), *Familiarity with Animated Characters* ($F(1,38)=0.01$, $p=0.93$), or in their *ages* ($F(1,38)=0.07$, $p=0.80$).

Therefore, any *Frame Rate* differences are not confounded by participant differences between groups.

Participants' mean score for the *Emotional Intelligence Questionnaire* was 104.13 out of 150 (SD=13.32). An average of 3.9 out of 7 (SD=2.38) was reported on *Experience in Video Games* and 4.9 out of 7 (SD=1.90) on *Familiarity with Animated Characters.*

| Anger | Disgust | Fear | Sadness | Shame | Happiness | Excitement | Pride | Relief | Surprise | Neutral |
|---|---|---|---|---|---|---|---|---|---|---|
| 97.5% | 95% | 85% | 85% | 75% | 95% | 90% | 100% | 60% | 80% | 52.5% |

**Table 5.1.** Percentage of participants that identified correctly the emotion at least once (chance level would be 24% for each emotion and 12% for neutral)

Prior to the full analysis, it was necessary to determine whether it was possible for participants to correctly identify each emotion from watching emotional body language alone (i.e. without facial or voice display). Recognition rates were above chance level (Table 5.1), although they varied between emotions (from min 60% for relief to max 100% for pride, chance level would be $1-(1/22)^6=24\%$).

## 5.2. Overview of the Experimental Effects

A high level analysis of the experimental conditions was carried out with Repeated Measures ANOVAs; 3 *Character Type* x 2 *Action Style* x 2 *Frame Rate*.

Character Type had no effect on Correct Identification ($F(1,38)=0.82$, $p=0.45$, Partial $\eta^2=0.02$) but it did have a significant effect on Strength ($F(1,38)=24.40$, $p<0.01$), Partial $\eta^2=0.39$), on Believability ($F(1,38)=16.84$, $p<0.01$, Partial $\eta^2=0.31$) and on Naturalness ($F(1,38)=15.66$, $p<0.01$, Partial $\eta^2=0.29$).

*Action Style* had a significant effect on *Correct Identification* ($F(1,38)=76.97$, $p<0.01$, Partial $\eta^2=0.67$), on *Strength* ($F(1,38)=160.53$, $p<0.01$, Partial $\eta^2=0.81$), on *Believability* ($F(1,38)=7.30$, $p<0.01$, Partial $\eta^2=0.16$) but not on *Naturalness* ($F(1,38)=0.46$, $p=0.50$, Partial $\eta^2=0.01$).

*Frame Rate*, had no effect on *Correct Identification* ($F(1,38)=0.06$, $p=0.80$, Partial $\eta^2=0.00$), on *Strength* ($F(1,38)=1.38$, $p=0.25$ Partial $\eta^2=0.03$), on *Believability* ($F(1,38)=0.12$, $p=0.74$, Partial $\eta^2=0.00$), nor on *Naturalness* ($F(1,38)= 0.21$, $p=0.65$ Partial $\eta^2=0.01$).

As suggested by the hypotheses (Section 4.2.3), the results showed significant interactions between *Character Type* and *Action Style* for *Strength* ($F(1,38)= 4.41$, $p<0.02$, Partial $\eta^2=0.10$) and for *Naturalness* ($F(1,38)= 3.40$, $p<0.05$, Partial $\eta^2=0.08$). Although it did not reach significance, there was a similar trend for *Believability* ($F(1,38)=2.64$, $p=0.08$, Partial $\eta^2=0.08$). However, there was no interaction between *Character Type* and *Action Style* for *Correct Identification* ($F(1,38)= 0.20$, $p=0.82$, Partial $\eta^2=0.01$). Moreover, there was no interaction between *Frame Rate* and any of the other independent variables.

Although, the combined results demonstrated an expected interaction, the recognition rate varied greatly depending on the emotions displayed (Table 5.1). As a comparable effect may have occurred for the other dependent variables, it would be prudent to consider the effect of *Frame Rate*, *Character Type* and *Action Style* on *Correct Identification*, *Strength*, *Believability* and on *Naturalness* for each Emotion Displayed. This subsequently led to a refinement of the 3 *Character Type* x 2 *Action Style* x 2 *Frame Rate* Repeated Measures ANOVAs' results (see Section 5.4).

# 5.3. Effects of Frame Rate

Frame rate was used to deteriorate the emotional body language displayed (Section 4.2.2). However, the overall trends suggest that *Frame Rate* had no effect on the dependent variables. A detailed analysis was still conducted to consider possible variations of the effect of *Frame Rate* on the dependent variables for the different emotions.

## 5.3.1. Effect of Frame Rate on Correct Identification

| Emotions | Effect of *Frame Rate* on the *Correct Identification* of the Ordinary Displays | Effect of *Frame Rate* on the *Correct Identification* of the Stylised Displays |
|---|---|---|
| Anger | $F(1,38)=0.28$, $p=0.60$, Partial $\eta^2=0.01$ | $F(1,38)=0.78$, $p=0.38$, Partial $\eta^2=0.02$ |
| Disgust | $F(1,38)=0.29$, $p=0.59$, Partial $\eta^2=0.01$ | $F(1,38)=2.80$, $p=0.10$, Partial $\eta^2=0.07$ |
| Fear | $F(1,38)=1.27$, $p=0.27$, Partial $\eta^2=0.03$ | $F(1,38)=0.46$, $p=0.50$, Partial $\eta^2=0.01$ |
| Sadness | $F(1,38)=0.95$, $p=0.34$, Partial $\eta^2=0.02$ | $F(1,38)=1.33$, $p=0.26$, Partial $\eta^2=0.03$ |
| Shame | $F(1,38)=0.00$, $p=1.00$, Partial $\eta^2=0.00$ | $F(1,38)=0.00$, $p=1.00$, Partial $\eta^2=0.00$ |
| Happiness | $F(1,38)=0.06$, $p=0.81$, Partial $\eta^2=0.00$ | $F(1,38)=1.89$, $p=0.18$, Partial $\eta^2=0.05$ |
| Excitement | $F(1,38)=0.14$, $p=0.71$, Partial $\eta^2=0.00$ | $F(1,38)=0.84$, $p=0.43$, Partial $\eta^2=0.02$ |
| Pride | $F(1,38)=0.18$, $p=0.67$, Partial $\eta^2=0.01$ | $F(1,38)=0.39$, $p=0.54$, Partial $\eta^2=0.01$ |
| Relief | $F(1,38)=0.00$, $p=1.00$, Partial $\eta^2=0.00$ | $F(1,38)=0.00$, $p=1.00$, Partial $\eta^2=0.00$ |
| Surprise | $F(1,38)=0.28$, $p=0.60$, Partial $\eta^2=0.01$ | $F(1,38)=3.28$, $p=0.08$, Partial $\eta^2=0.08$ |
| Neutral | $F(1,38)=0.54$, $p=0.46$, Partial $\eta^2=0.01$ | |

**Table 5.2.** Results of the ANOVAs carried out on *Correct Identification* between *Frame Rate* for each Emotion Displayed (Shaded results are significant). Ordinary and Stylised displays are separate.

To investigate the effect of *Frame Rate* on *Correct Identification*, ANOVAs were carried out on the data for each emotional display (Table 5.2). There was no effect of *Frame Rate* on *Correct Identification* for any emotion (Table 5.2) and Partial $\eta^2$ from 0.01 to 0.08 which suggests that there was almost no difference between the two Frame Rate conditions.

Participants were thus similar in their performance of identifying the emotional body language displayed regardless of the *Frame Rate* condition they were in.

### 5.3.2. Effect of Frame Rate on Strength

| Emotions | Effect of *Frame Rate* on the *Strength* of the Ordinary Displays | Effect of *Frame Rate* on the *Strength* of the Stylised Displays |
|---|---|---|
| **Anger** | $F(1,38)=5.03$, $p<0.05$, Partial $\eta^2=0.12$ | $F(1,38)=0.51$, $p=0.48$, Partial $\eta^2=0.01$ |
| **Disgust** | $F(1,38)=1.96$, $p=0.17$, Partial $\eta^2=0.05$ | $F(1,38)=0.05$, $p=0.83$, Partial $\eta^2=0.00$ |
| **Fear** | $F(1,38)=0.38$, $p=0.54$, Partial $\eta^2=0.01$ | $F(1,38)=1.09$, $p=0.30$, Partial $\eta^2=0.03$ |
| **Sadness** | $F(1,38)=4.17$, $p<0.05$, Partial $\eta^2=0.10$ | $F(1,38)=0.57$, $p=0.45$, Partial $\eta^2=0.01$ |
| **Shame** | $F(1,38)=2.31$, $p=0.14$, Partial $\eta^2=0.06$ | $F(1,38)=0.01$, $p=0.94$, Partial $\eta^2=0.00$ |
| **Happiness** | $F(1,38)=0.39$, $p=0.53$, Partial $\eta^2=0.01$ | $F(1,38)=0.36$, $p=0.55$, Partial $\eta^2=0.01$ |
| **Excitement** | $F(1,38)=0.64$, $p=0.43$, Partial $\eta^2=0.02$ | $F(1,38)=0.00$, $p=1.00$, Partial $\eta^2=0.00$ |
| **Pride** | $F(1,38)=0.98$, $p=0.33$, Partial $\eta^2=0.02$ | $F(1,38)=0.01$, $p=0.94$, Partial $\eta^2=0.00$ |
| **Relief** | $F(1,38)=1.70$, $p=0.20$, Partial $\eta^2=0.04$ | $F(1,38)=0.04$, $p=0.85$, Partial $\eta^2=0.00$ |
| **Surprise** | $F(1,38)=1.98$, $p=0.17$, Partial $\eta^2=0.05$ | $F(1,38)=0.23$, $p=0.63$, Partial $\eta^2=0.01$ |
| **Neutral** | $F(1,38)=0.86$, $p=0.36$, Partial $\eta^2=0.02$ | |

**Table 5.3.** Results of the ANOVAs carried out on *Strength* between *Frame Rate* for each Emotion Displayed (Shaded results are significant). Ordinary and Stylised displays are separate.

To investigate the effect of *Frame Rate* on *Strength*, ANOVAs were carried out on the data for each Emotion Displayed (Table 5.3). With the exception of "Anger Ordinary" and "Sadness Ordinary", there was no effect of *Frame Rate* on *Strength* (Table 5.3), with Partial $\eta^2$ ranged from 0.01 to 0.12.

Thus, the results suggest that participants' perception of the emotions' *Strength* was not broadly affected by the change in *Frame Rate*.

## 5.3.3. Effect of Frame Rate on Believability

| Emotions | Effect of *Frame Rate* on the *Believability* of the Ordinary Displays | Effect of *Frame Rate* on the *Believability* of the Stylised Displays |
|---|---|---|
| **Anger** | $F(1,38)=1.21$, $p=0.28$, Partial $\eta^2=0.03$ | $F(1,38)=0.07$, $p=0.79$, Partial $\eta^2=0.00$ |
| **Disgust** | $F(1,38)=0.35$, $p=0.56$, Partial $\eta^2=0.01$ | $F(1,38)=0.02$, $p=0.89$, Partial $\eta^2=0.00$ |
| **Fear** | $F(1,38)=0.62$, $p=0.43$, Partial $\eta^2=0.02$ | $F(1,38)=0.29$, $p=0.59$, Partial $\eta^2=0.01$ |
| **Sadness** | $F(1,38)=0.05$, $p=0.81$, Partial $\eta^2=0.00$ | $F(1,38)=0.32$, $p=0.57$, Partial $\eta^2=0.01$ |
| **Shame** | $F(1,38)=1.11$, $p=0.30$, Partial $\eta^2=0.03$ | $F(1,38)=0.00$, $p=1.00$, Partial $\eta^2=0.00$ |
| **Happiness** | $F(1,38)=1.23$, $p=0.27$, Partial $\eta^2=0.03$ | $F(1,38)=1.07$, $p=0.31$, Partial $\eta^2=0.03$ |
| **Excitement** | $F(1,38)=0.02$, $p=0.89$, Partial $\eta^2=0.00$ | $F(1,38)=0.02$, $p=0.87$, Partial $\eta^2=0.00$ |
| **Pride** | $F(1,38)=0.09$, $p=0.77$, Partial $\eta^2=0.00$ | $F(1,38)=0.04$, $p=0.83$, Partial $\eta^2=0.00$ |
| **Relief** | $F(1,38)=1.23$, $p=0.27$, Partial $\eta^2=0.03$ | $F(1,38)=0.04$, $p=0.83$, Partial $\eta^2=0.00$ |
| **Surprise** | $F(1,38)=0.56$, $p=0.46$, Partial $\eta^2=0.01$ | $F(1,38)=0.11$, $p=0.74$, Partial $\eta^2=0.00$ |
| **Neutral** | $F(1,38)=0.39$, $p=0.54$, Partial $\eta^2=0.01$ | |

**Table 5.4.** Results of the ANOVAs carried out on *Believability* between *Frame Rate* for each Emotion Displayed (Shaded results are significant). Ordinary and Stylised displays are separate.

To investigate the effect of *Frame Rate* on *Believability*, ANOVAs were carried out on the data for each *Emotion Displayed* (Table 5.4). There was no effect of *Frame Rate* on *Believability* (Table 5.4) with the Partial $\eta^2$ ranging from 0.01 to 0.03.

Participants rated the believability of each emotional display equally, regardless of the frame rate condition they were in.

### 5.3.4. Effect of Frame Rate on Naturalness

| Emotions | Effect of *Frame Rate* on the *Naturalness* of the Ordinary Displays | Effect of *Frame Rate* on the *Naturalness* of the Stylised Displays |
|---|---|---|
| **Anger** | $F(1,38)=0.38$, $p=0.54$, Partial $\eta^2=0.01$ | $F(1,38)=0.03$, $p=0.86$, Partial $\eta^2=0.00$ |
| **Disgust** | $F(1,38)=0.01$, $p=0.93$, Partial $\eta^2=0.00$ | $F(1,38)=0.31$, $p=0.58$, Partial $\eta^2=0.01$ |
| **Fear** | $F(1,38)=1.13$, $p=0.29$, Partial $\eta^2=0.03$ | $F(1,38)=0.10$, $p=0.76$, Partial $\eta^2=0.00$ |
| **Sadness** | $F(1,38)=0.21$, $p=0.65$, Partial $\eta^2=0.01$ | $F(1,38)=0.01$, $p=0.91$, Partial $\eta^2=0.00$ |
| **Shame** | $F(1,38)=1.70$, $p=0.20$, Partial $\eta^2=0.04$ | $F(1,38)=0.08$, $p=0.77$, Partial $\eta^2=0.00$ |
| **Happiness** | $F(1,38)=0.08$, $p=0.77$, Partial $\eta^2=0.00$ | $F(1,38)=0.08$, $p=0.78$, Partial $\eta^2=0.00$ |
| **Excitement** | $F(1,38)=0.48$, $p=0.49$, Partial $\eta^2=0.01$ | $F(1,38)=0.13$, $p=0.72$, Partial $\eta^2=0.00$ |
| **Pride** | $F(1,38)=1.20$, $p=0.28$, Partial $\eta^2=0.03$ | $F(1,38)=0.09$, $p=0.76$, Partial $\eta^2=0.00$ |
| **Relief** | $F(1,38)=3.82$, $p=0.06$, Partial $\eta^2=0.09$ | $F(1,38)=0.15$, $p=0.70$, Partial $\eta^2=0.00$ |
| **Surprise** | $F(1,38)=2.52$, $p=0.12$, Partial $\eta^2=0.06$ | $F(1,38)=0.00$, $p=0.96$, Partial $\eta^2=0.00$ |
| **Neutral** | $F(1,38)=0.02$, $p=0.89$, Partial $\eta^2=0.00$ | |

**Table 5.5.** Results of the ANOVAs carried out on *Naturalness* between *Frame Rate* for each Emotion Displayed (Shaded results are significant). Ordinary and Stylised displays are separate.

To investigate the effect of *Frame Rate* on *Naturalness*, ANOVAs were carried on the data for each Emotion Displayed (Table 5.5). The results show that there was no effect of *Frame Rate* on *Naturalness* (Table 5.5), with the Partial $\eta^2$ ranging from 0.00 to 0.09.

Participants thus rated the naturalness of each emotional display equally, regardless of the frame rate condition they were in.

### 5.3.5. Summary of the effects of *Frame Rate*

*Frame Rate* had no significant effect on the different dependent variables. Therefore, in order to get a larger sample size and to remove the between group constraint on the statistical tests, the two sets of data will be combined for the rest of the analysis, effectively eliminating *Frame Rate* as a condition. Thus, the overall results will need to be reassessed with the analysis for *Character Type* and *Action Style* combining the two *Frame Rate* conditions.

## 5.4. Reinvestigation of the Effects of Character Type on Correct Identification, Strength, Believability and Naturalness

The effects of *Character Type* were reassessed with 3 (*Character Type*) x 2 (*Action Style*) Repeated Measures ANOVAs on each dependent variable. This will reassess the general patterns (reported in Section 5.2) after eliminating the between subjects *Frame Rate* condition (Note: the analysis of the interactions between *Character Type* and *Action Style* from these tests is reported in Section 5.7).

As in the original analysis (see Section 5.2), *Character Type* still had no effect on *Correct Identification* (F(1,38)=0.84, p=0.44, Partial $\eta^2$=0.02).

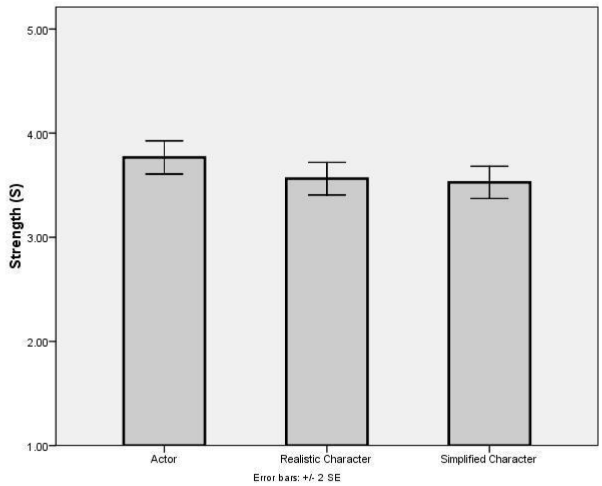


**Figure 5.1.** The overall *Strength* by *Character Type* (Means + 2*Standard Error).

*Character Type* did still have a significant effect on *Strength* (F(1,38)=23.99, p<0.01, Partial $\eta^2$=0.38). Post Hoc analysis (Least Significant Difference) showed that overall, the emotions were perceived as stronger when displayed by the Actor than by the Realistic Character (p<0.01) or by the Simplified Character (p<0.01). However, they were perceived as similar in terms of *Strength* when displayed by the Realistic Character and the Simplified Character (p=0.40) (Figure 5.1).

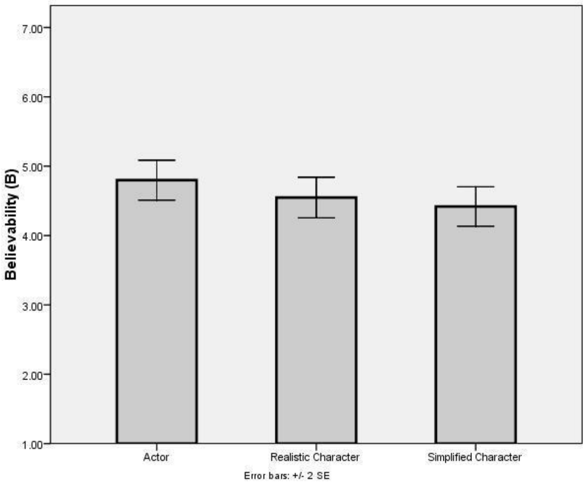


**Figure 5.2.** The overall *Believability* by *Character Type* (Means + 2*Standard Error).

*Character Type* also still had a significant effect on *Believability* ($F(1,38)=17.22$, $p<0.01$, Partial $\eta^2=0.31$). Post Hoc analysis (Least Significant Difference) showed that, overall, the emotions were perceived as more *Believable* when displayed by the Actor than by the Realistic Character ($p<0.01$) or by the Simplified Character ($p<0.01$). Moreover, they were perceived as more *Believable* when displayed by the Realistic Character than by the Simplified Character ($p<0.01$) (Figure 5.2).

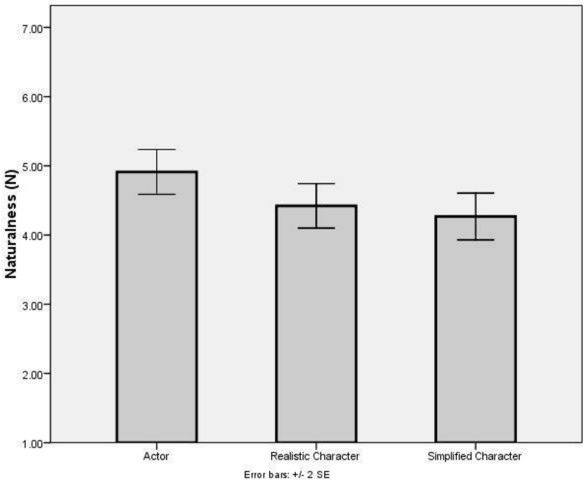


**Figure 5.3.** The overall *Naturalness* by *Character Type* (Means + 2*Standard Error).

*Character Type* again still had a significant effect on *Naturalness* ($F(1,38)=15.82$, $p<0.01$, Partial $\eta^2=0.29$). Post Hoc analysis (Least Significant Difference) showed that, overall, the emotions were perceived as more *Natural* when displayed by the Actor than by the Realistic Character ($p<0.01$) or by the Simplified Character ($p<0.01$). Moreover, they were perceived as more *Natural* when displayed by the Realistic Character than by the Simplified Character ($p<0.05$) (Figure 5.3).

The general patterns from the reinvestigation indicate that, overall, participants were equally good at interpreting the emotions displayed by the actor, the realistic character and the simplified character. However, the character displaying the emotions did have an effect on *Strength* (Figure 5.1), on *Believability* (Figure 5.2) and on *Naturalness* (Figure 5.3), although this is the same as in the original analysis that included *Frame Rate*.

Thus, given the similarity between these combined results and the original ones (Section 5.2), these combined results replace the original ones. The detailed analysis of the effect of *Character Type* on *Correct Identification*, *Strength*, *Believability* and on *Naturalness* for each *Emotion Displayed* can thus now be conducted on this combined set of data.

# 5.5. Effect of Character Type

## 5.5.1. Effect of Character Type on Correct Identification

| Emotions | Effect of *Character Type* on the *Correct Identification* for the Ordinary Displays | Effect of *Character Type* on the *Correct Identification* of the Stylised Displays |
|---|---|---|
| Anger | $F(1,38)=2.00$, $p=0.15$, Partial $\eta^2=0.05$ | $F(1,38)=0.32$, $p=0.71$, Partial $\eta^2=0.01$ |
| Disgust | $F(1,38)=1.70$, $p=0.19$ , Partial $\eta^2=0.04$ | **$F(1,38)=3.84$, $p<0.05$, Partial $\eta^2=0.09$. AC=RC ($p=0.29$), AC>SC ($p<0.05$), RC=SC ($p=0.06$)** |
| Fear | **$F(1,38)=8.41$, $p<0.01$, Partial $\eta^2=0.18$. AC>RC ($p<0.01$), AC>SC ($p<0.01$), RC=SC ($p=0.34$)** | $F(1,38)=2.11$, $p=0.13$, Partial $\eta^2=0.05$ |
| Sadness | $F(1,38)=0.48$, $p=0.62$, Partial $\eta^2=0.01$ | $F(1,38)=1.60$, $p=0.21$, Partial $\eta^2=0.04$ |
| Shame | **$F(1,38)=5.38$, $p<0.01$, Partial $\eta^2=0.12$. AC>RC ($p<0.05$), AC>SC ($p<0.01$), RC=SC ($p=0.24$)** | $F(1,38)=1.50$, $p=0.23$, Partial $\eta^2=0.04$ |
| Happiness | **$F(1,38)=4.13$, $p<0.05$, Partial $\eta^2=0.10$. AC=RC ($p=0.83$), AC<SC ($p<0.05$), RC<SC ($p<0.05$)** | $F(1,38)=1.21$, $p=0.31$, Partial $\eta^2=0.03$ |
| Excitement | $F(1,38)=0.81$, $p=0.45$, Partial $\eta^2=0.02$ | **$F(1,38)=3.49$, $p<0.05$, Partial $\eta^2=0.08$. AC<RC ($p<0.05$), AC=SC ($p=0.39$), RC>SC ($p<0.05$)** |
| Pride | **$F(1,38)=3.38$, $p<0.05$, Partial $\eta^2=0.08$. AC=RC ($p=0.57$), AC>SC ($p<0.01$), RC=SC ($p=0.09$)** | $F(1,38)=1.52$, $p=0.22$, Partial $\eta^2=0.04$ |
| Relief | **$F(1,38)=7.92$, $p<0.01$, Partial $\eta^2=0.17$. AC=RC ($p=0.08$), AC>SC ($p<0.01$), RC>SC ($p<0.05$)** | $F(1,38)=0.39$, $p=0.67$, Partial $\eta^2=0.01$ |
| Surprise | $F(1,38)=1.28$, $p=0.29$, Partial $\eta^2=0.03$ | $F(1,38)=0.59$, $p=0.53$, Partial $\eta^2=0.01$ |
| Neutral | $F(1,38)=0.98$, $p=0.38$, Partial $\eta^2=0.02$ | |

**Table 5.6.** Results of the Repeated Measures ANOVAs carried out on *Correct Identification* between *Character Type* (AC=Actor, RC=Realistic Character, SC=Simplified Character) for each Emotion Displayed (Shaded results are significant). Ordinary and Stylised displays are separated.

To assess the effect of *Character Type* for each Emotion Displayed, repeated measures ANOVAs were carried out on the data. For fourteen of the twenty-one results there was no effect of *Character Type* on *Correct identification* (Table 5.6).

However, there were seven results for which an effect was found (Table 5.6). For these, post-hoc analysis (Least Significant Difference) was conducted to investigate which characters differed and in what direction (Table 5.6). Results showed no consistent direction or effect. For example, for “Fear Ordinary”, the participants were better at interpreting the Actor than the Realistic Character or the Simplified one. However, for “Excitement Stylised”, participants were better in interpreting the Realistic Character than the Actor or the Simplified Character.

The lack of significant differences in the *Correct Identification* for most of the displays across the different *Character Types* could indicate that interpretation of emotional body language did not depend on the presentation condition. However, to confirm this and eliminate the possibility that it was due to a lack of statistical power or to randomness in the interpretation, a psychological space of the interpretation was generated using a multidimensional scaling procedure (PROXSCAL).

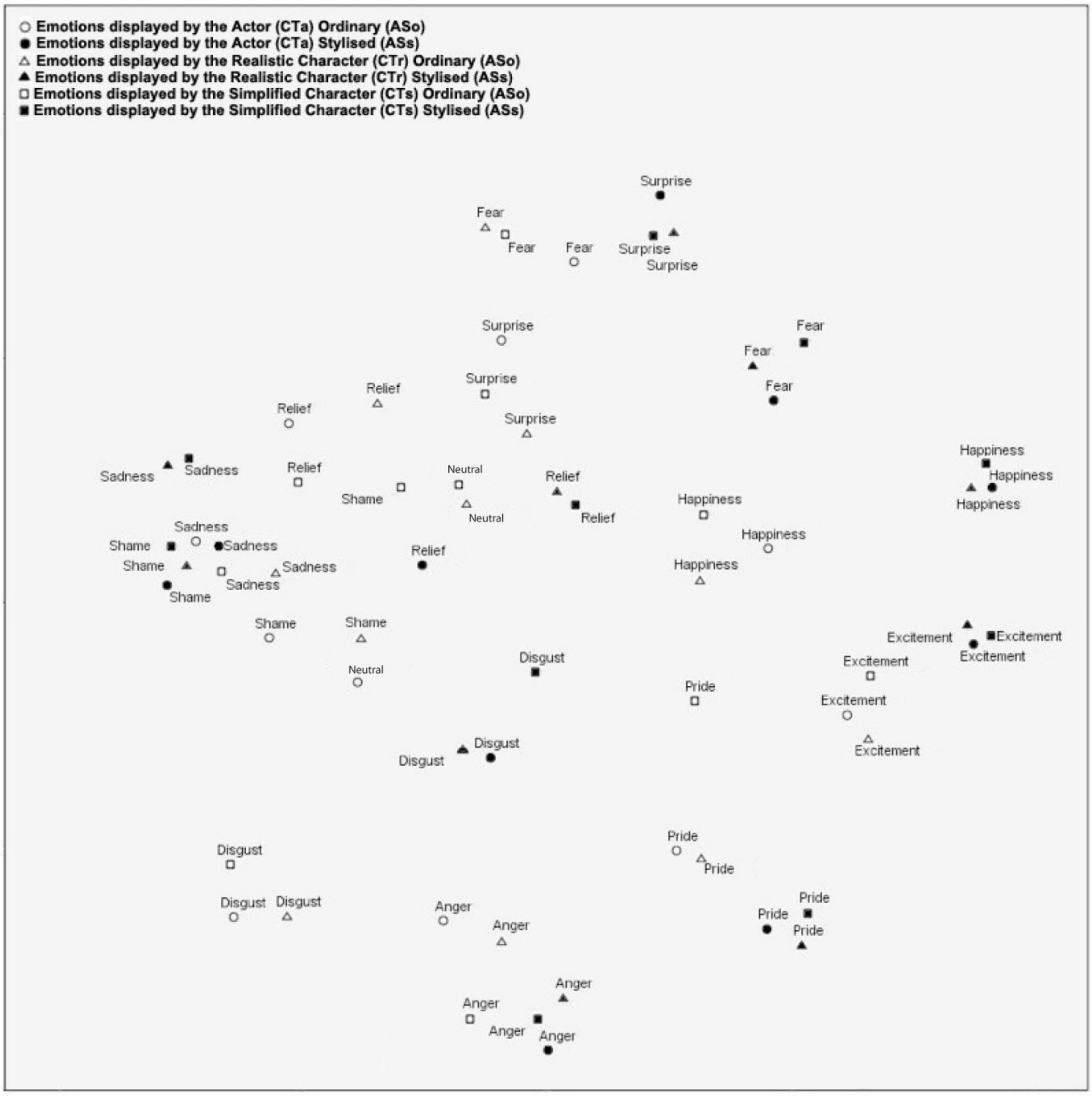


**Figure 5.4.** The psychological space of interpretation of emotions display obtained by Multi Dimensional Scaling.

Multidimensional scaling is a systematic procedure for obtaining a spatial representation consisting of a geometric configuration of points (Kruskal & Wish, 1978). Each point in the configuration (Figure 5.4) corresponds to one of the responses for the 63 presented emotion clips (10 *Emotion Displayed* x 2 *Action Style* (Natural and Stylised) + 1 neutral state) x 3 *Character Types*). Multidimensional scaling uses a measure of proximity between objects as input (each videos clip presented in this case). The proximities between presented emotion clips were generated by counting the percentage of interpretations each clip had in common with the others. Clips that were interpreted similarly had smaller distances from one another than those that were not. Presented emotion clips were labeled by intended emotion (rather than by interpretations). The resultant solution is a unique two-dimensional psychological space (Figure 5.4).

The clustering of videos clips in the psychological space (Figure 5.4) show that the intended emotions cluster together independently from the character presented. This suggests that the lack of general significant differences (see results in Section 5.4) is not due to randomness in the interpretation or to a lack of statistical power that could have made differences undetectable. It rather suggests that the videos clips were interpreted independently from the character (Figure 5.4).

Overall, although it was not the case for all the emotions displayed, the detailed analysis suggests that *Character Type* did not affect participants when identifying the Emotions Displayed.

### 5.5.2. Effect of Character Type on Strength

| Emotions | Effect of *Character Type* on the *Strength* for the Ordinary Displays | Effect of *Character Type* on the *Strength* for the Stylised Displays |
|---|---|---|
| Anger | F(1,38)=1.61, p=0.21, Partial $\eta^2$=0.04 | **F(1,38)=5.84, p<0.01, Partial $\eta^2$=0.13. AC>RC (p<0.05), AC>SC (P<0.01), RC=SC (p=0.23)** |
| Disgust | F(1,38)==2.52, P=0.09, Partial $\eta^2$=0.06 | F(1,38)==0.69, P=0.49, Partial $\eta^2$=0.02 |
| Fear | F(1,38)==0.46, P=0.63, Partial $\eta^2$=0.01 | F(1,38)==2.84, P=0.06, Partial $\eta^2$=0.07 |
| Sadness | **F(1,38)==4.45, P<0.05, Partial $\eta^2$=0.10. AC=RC (p<0.07), AC>SC (P<0.05), RC=SC (p=0.20)** | F(1,38)==0.90, P=0.41, Partial $\eta^2$=0.02 |
| Shame | **F(1,38)==14.08, P<0.01, Partial $\eta^2$=0.27. AC>RC (p<0.01), AC>SC (P<0.01), RC>SC (p<0.05)** | **F(1,38)==5.74, P<0.01, Partial $\eta^2$=0.30. AC>RC (p<0.01), AC>SC (P<0.01), RC=SC (p=1.00)** |
| Happiness | F(1,38)==1.62, P=0.20, Partial $\eta^2$=0.04 | F(1,38)==0.44, P=0.64, Partial $\eta^2$=0.01 |
| Excitement | F(1,38)==1.54, P=0.22, Partial $\eta^2$=0.04 | F(1,38)==1.88, P=0.17, Partial $\eta^2$=0.05 |
| Pride | **F(1,38)==8.97, P<0.01, Partial $\eta^2$=0.19. AC=RC (p=0.50), AC>SC (P<0.01), RC>SC (p<0.01)** | F(1,38)==2.62, P=0.08, Partial $\eta^2$=0.06 |
| Relief | **F(1,38)==7.89, P<0.01, Partial $\eta^2$=0.17. AC>RC (p<0.01), AC=SC (P=0.06), RC=SC (p=0.07)** | F(1,38)==0.73, P=0.48, Partial $\eta^2$=0.02 |
| Surprise | F(1,38)==1.23, P=0.30, Partial $\eta^2$=0.03 | F(1,38)==0.66, P=0.52, Partial $\eta^2$=0.02 |
| Neutral | F(1,38)==0.02, P=0.98, Partial $\eta^2$=0.00 | |

**Table 5.7.** Results of the Repeated Measures ANOVAs carried out on *Strength* between *Character Type* (AC=Actor, RC=Realistic Character, SC=Simplified Character) for each Emotion Displayed (Shaded results are significant). Ordinary and Stylised displays are separates.

To assess the effect of *Character Type* on *Strength* for each *Emotion Displayed*, repeated measures ANOVAs were carried out on the data. For fifteen of the twenty-one, there was no effect of *Character Type* on *Strength* (Table 5.7).

However, there were six Emotions Displayed for which an effect was found (Table 5.7). For these six displays, post-hoc analysis (Least Significant Difference) was conducted to investigate how characters differed and in what direction (Table 5.7). Although the pattern was not identical, it generally showed that the *Emotion Displayed* were perceived as stronger (*Strength*) when displayed by the Actor than by the Realistic Character or by the Simplified Character (Table 5.7). The emotions displayed by the Realistic Character and by the Simplified Character were generally perceived similarly in term of *Strength* (Table 6.7). Although, if evident, the Emotions Displayed by the Realistic Character are perceived as stronger than when displayed by the Simplified Character. These six displays are likely to be responsible for the overall differences that were found in *Strength* across *Character Type* (see Section 5.4).

The detailed analysis suggests that, overall, participants were not affected by the *Character Type* when interpreting the emotional *Strength* and that the general difference is rather due to 6 Emotions Displayed for which an effect similar to the overall results (see Section 5.4) was found.

### 5.5.3. Effect of Character Type on Believability

| Emotions | Effect of *Character Type* on the *Believability* for the Ordinary Displays | Effect of *Character Type* on the *Believability* for the Stylised Displays |
|---|---|---|
| Anger | $F(1,38)=1.43$, $P=0.25$, Partial $\eta^2=0.04$ | $F(1,38)=2.40$, $P=0.10$, Partial $\eta^2=0.06$ |
| Disgust | **$F(1,38)=7.84$, $P<0.01$, Partial $\eta^2=0.17$. AC>RC ($p<0.05$), AC>SC ($P<0.01$), RC=SC ($p=0.11$)** | $F(1,38)=0.09$, $P=0.90$, Partial $\eta^2=0.00$ |
| Fear | $F(1,38)=0.11$, $P=0.89$, Partial $\eta^2=0.00$ | $F(1,38)=2.44$, $P=0.09$, Partial $\eta^2=0.06$ |
| Sadness | **$F(1,38)=5.21$, $P<0.01$, Partial $\eta^2=0.12$. AC>RC ($p<0.01$), AC>SC ($P<0.05$), RC=SC ($p=0.33$)** | $F(1,38)=0.57$, $P=0.57$, Partial $\eta^2=0.01$ |
| Shame | **$F(1,38)=8.75$, $P<0.01$, Partial $\eta^2=0.19$. AC>RC ($p<0.05$), AC>SC ($P<0.01$), RC=SC ($p=0.07$)** | **$F(1,38)=4.57$, $P<0.05$, Partial $\eta^2=0.11$. AC>RC ($p<0.01$), AC>SC ($P<0.05$), RC=SC ($p=0.84$)** |
| Happiness | $F(1,38)=1.86$, $P=0.16$, Partial $\eta^2=0.05$ | $F(1,38)=0.14$, $P=0.87$, Partial $\eta^2=0.00$ |
| Excitement | $F(1,38)=1.22$, $P=0.30$, Partial $\eta^2=0.03$ | $F(1,38)=0.21$, $P=0.76$, Partial $\eta^2=0.01$ |
| Pride | **$F(1,38)=6.42$, $P<0.01$, Partial $\eta^2=0.14$. AC=RC ($p=0.06$), AC>SC ($P<0.01$), RC=SC ($p=0.12$)** | $F(1,38)=0.59$, $P=0.56$, Partial $\eta^2=0.02$ |
| Relief | $F(1,38)=2.88$, $P=0.06$, Partial $\eta^2=0.07$ | $F(1,38)=3.16$, $P=0.05$, Partial $\eta^2=0.08$ |
| Surprise | $F(1,38)=0.70$, $P=0.50$, Partial $\eta^2=0.02$ | **$F(1,38)=3.15$, $P<0.05$, Partial $\eta^2=0.08$ AC=RC ($p=0.24$), AC>SC ($P<0.05$), RC=SC ($p=0.21$)** |
| Neutral | $F(1,38)=2.86$, $P=0.07$, Partial $\eta^2=0.07$ | |

**Table 5.8.** Results of the repeated measures ANOVA and post-hoc tests carried on *Believability* between *Character Type* (AC= Actor, RC= Realistic Character, SC= Simplified Character). For Ordinary and Stylised displays separately.

To assess the effect of *Character Type* on *Believability* (defined as the extent to which a character was feeling the emotion), repeated measures ANOVAs were carried out on the data (Table 5.8). For six of the twenty-one Emotions Displayed there was an effect. Post-hocs tests (Least Significant Difference) showed that, when an effect was present, the Actor was more or equally believable than the Realistic Character, which was equally believable to the Simplified Character (Table 5.8). These results are slightly different from the overall results (see Section 5.4) for which the Realistic Character was found more Believable than the Simplified Character. This suggests that the overall results are not the consequence of a few specific Emotions Displayed but are rather due to an overall difference. This is further suggested by the trends present in four of the twenty-one displays ($p<0.1$) (Table 5.8). Thus, the detailed analysis confirmed the overall effect of *Character Type* on *Believability* that was found (see Section 5.4).

Although it was not the case for all the displays, participants found that the emotional body language was more believable when displayed by the Actor than when displayed by the Realistic Character and it was found less believable when displayed by the Simplified Character than when displayed by the two others.

### 5.5.4. Effect of Character Type on Naturalness

| Emotions | Effect of *Character Type* on the *Strength* for the Ordinary Displays | Effect of *Character Type* on the *Strength* for the Stylised Displays |
|---|---|---|
| Anger | **$F(1,38)=6.75$, $P<0.01$, Partial $\eta^2=0.15$.<br>AC=RC ($p=0.09$), AC>SC ($P<0.01$), RC>SC ($p<0.05$)** | **$F(1,38)=4.82$, $P<0.05$, Partial $\eta^2=0.11$.<br>AC>RC ($p<0.05$), AC>SC ($P<0.01$), RC=SC ($p=1.00$)** |
| Disgust | **$F(1,38)=9.18$, $P<0.01$, Partial $\eta^2=0.20$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC=SC ($p=0.11$)** | **$F(1,38)=9.14$, $P<0.01$, Partial $\eta^2=0.20$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC=SC ($p=0.59$)** |
| Fear | $F(1,38)=0.53$, $P=0.59$, Partial $\eta^2=0.01$ | **$F(1,38)=5.96$, $P<0.01$, Partial $\eta^2=0.13$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC=SC ($p=0.82$)** |
| Sadness | **$F(1,38)=9.10$, $P<0.01$, Partial $\eta^2=0.19$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC=SC ($p=0.60$)** | $F(1,38)=1.27$, $P=0.29$, Partial $\eta^2=0.03$ |
| Shame | **$F(1,38)=11.48$, $P<0.01$, Partial $\eta^2=0.23$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC>SC ($p<0.01$)** | **$F(1,38)=5.68$, $P<0.01$, Partial $\eta^2=0.13$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.05$), RC=SC ($p=0.48$)** |
| Happiness | **$F(1,38)=4.56$, $P<0.05$, Partial $\eta^2=0.11$.<br>AC=RC ($p=0.10$), AC>SC ($P<0.05$), RC=SC ($p<0.14$)** | $F(1,38)=1.49$, $P=0.23$, Partial $\eta^2=0.04$ |
| Excitement | $F(1,38)=2.03$, $P=0.14$, Partial $\eta^2=0.05$ | $F(1,38)=3.72$, $P=0.10$, Partial $\eta^2=0.06$ |
| Pride | **$F(1,38)=12.16$, $P<0.01$, Partial $\eta^2=0.24$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC>SC ($p<0.05$)** | $F(1,38)=3.14$, $P=0.06$, Partial $\eta^2=0.08$ |
| Relief | **$F(1,38)=8.56$, $P<0.01$, Partial $\eta^2=0.18$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC=SC ($p=0.46$)** | $F(1,38)=2.36$, $P=0.10$, Partial $\eta^2=0.06$ |
| Surprise | $F(1,38)=1.94$, $P=0.15$, Partial $\eta^2=0.05$ | **$F(1,38)=5.48$, $P<0.01$, Partial $\eta^2=0.12$.<br>AC>RC ($p<0.05$), AC>SC ($P<0.01$), RC=SC ($p=0.42$)** |
| Neutral | **$F(1,38)=8.63$, $P<0.01$, Partial $\eta^2=0.19$.<br>AC>RC ($p<0.01$), AC>SC ($P<0.01$), RC=SC ($p=0.92$)** | |

**Table 5.9.** Results of the repeated measures ANOVA and post-hoc tests carried on *Naturalness* between *Character Type* (AC= Actor, RC= Realistic Character, SC= Simplified Character). For Ordinary and Stylised displays separately.

To assess the effect of *Character Type* on *Naturalness* (defined as the quality of the character's movements), repeated measures ANOVAs were carried out on the data (Table 5.9). For thirteen of the twenty-one Emotions Displayed there was an effect (Table 5.9). Post-hocs tests Least Significant Differences) showed that the differences between characters were similar for most of the effects found (Table 5.9) with the Actor found more natural than the Realistic Character or the Simplified Character and the Realistic Character was found more natural than the Simplified Character. These results confirm the overall effect of *Character Type* on *Naturalness* that was found (see Section 5.4).

Thus, participants found that the emotional body language was more natural when displayed by the Actor than when displayed by the Realistic Character and it was found less natural when displayed by the Simplified Character than when displayed by the two others.

### 5.5.5. Summary of the effects of *Character Type*

Overall, *Character Type* did not broadly affect the participants' identification of the emotional body language. There were, however, a few exceptions. For most of these, participants were better or equal in identifying the emotion when displayed by the Actor than when displayed by the Realistic Character or the Simplified Character. Surprisingly, there was one Emotion Displayed for which participants were significantly better in interpreting the Simplified Character than the Actor or the Realistic Character. Nevertheless, the results of the multidimensional scaling (see Section 5.5.1) suggest that these are exceptions and that participants interpreted the Emotions Displayed by the three characters similarly.

However, the overall results (see Section 5.4) showed a significant effect of *Character Type* on *Strength*, although, the detailed analysis suggested that this was due to a few exceptions as it was found that for most of the displays participants were not affected by the changes of *Character Type*.

The perceived *Believability* and *Naturalness* of the movements were also affected by the change in *Character Type* (see results reported in Section 5.5.3 and Section 5.5.4). Overall, the emotions displayed by the Actor were found more *Believable* and more *Natural* than the ones displayed by the Realistic Character, which were found more *Believable* and more *Natural* than the ones displayed by the Simplified Character.

## 5.6. Reinvestigation of the Effects of Action Style

The effects of *Action Style* were reassessed with 3 (*Character Type*) x 2 (*Action Style*) Repeated Measures ANOVAs on each dependent variable to determine any difference after eliminating the between subjects *Frame Rate* condition (The analysis of the effect of *Character Type* from these tests was reported in section 5.4).

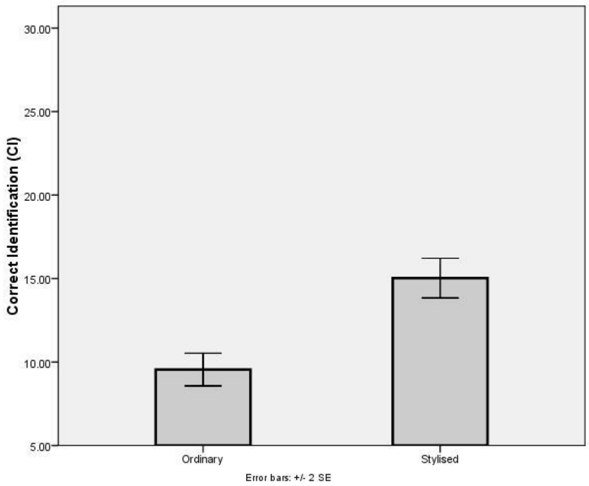


**Figure 5.5.** The Overall *Correct Identification* for the two styles (Means + 2*Standard Error)

*Action Style* did still have a significant effect on *Correct Identification* ($F(1,38)=78.73$, $p<0.01$, Partial $\eta^2=0.67$) (Figure 5.5), with the stylised display higher than the ordinary ones.

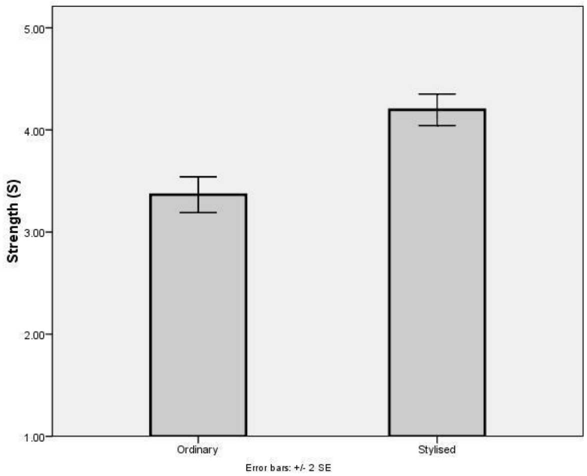


**Figure 5.6.** The Overall *Strength* for the two styles (Means + 2*Standard Error)

*Action Style* still had a significant effect on *Strength* ($F(1,38)=134,54$, $p<0.01$, Partial $\eta^2=0.77$) (Figure 5.6), with stylised displays higher than the ordinary ones.

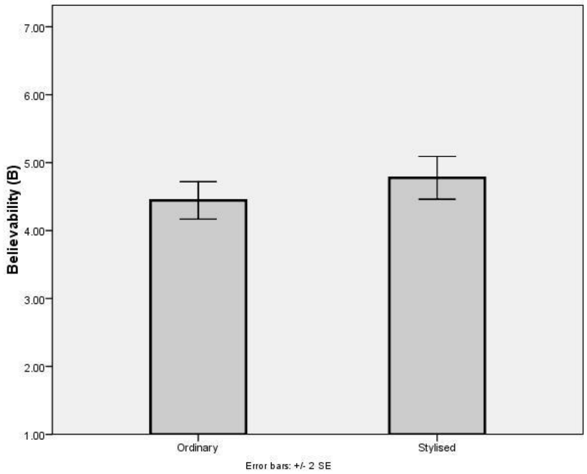


**Figure 5.7.** The Overall *Believability* for the two styles (Means + 2*Standard Error)

*Action Style* still had a significant effect on *Believability* ($F(1,38)=7.03$, $p<0.05$, Partial $\eta^2=0.15$) (Figure 5.7), with stylised displays higher than the ordinary ones.

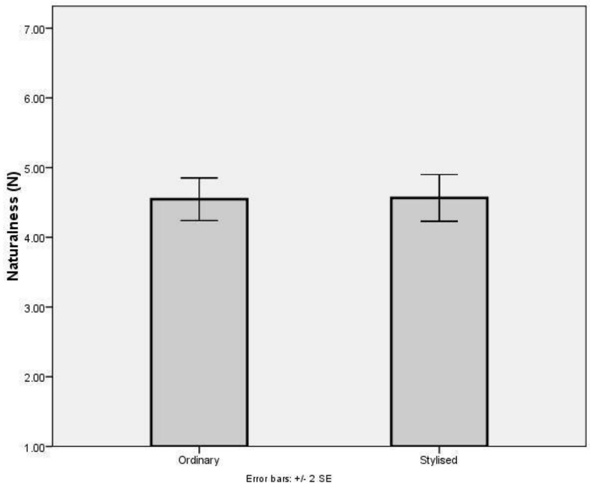


**Figure 5.8.** The Overall *Naturalness* for the two styles (Means + 2*Standard Error)

*Action Style* also still had no effect on *Naturalness* ($F(1,38)=0.42$, $p=0.52$, Partial $\eta^2=0.01$) (Figure 5.8), with stylised displays higher than the ordinary ones.

As in the original analysis (see Section 5.2), the patterns from the reinvestigation indicate that, overall, the Stylised displays were better identified, were found more emotional (*Strength*) and more *Believable*. There was no difference in *Naturalness*.

## 5.7. Interactions between Action Style and Character Type.

At this stage, the interactions between *Action Style* and *Character Type* were reinvestigated. There was still no interaction between *Action Style* and *Character Type* for *Correct Identification* ($F(1,38)=0.20$, $p=0.82$, Partial $\eta^2=0.01$).

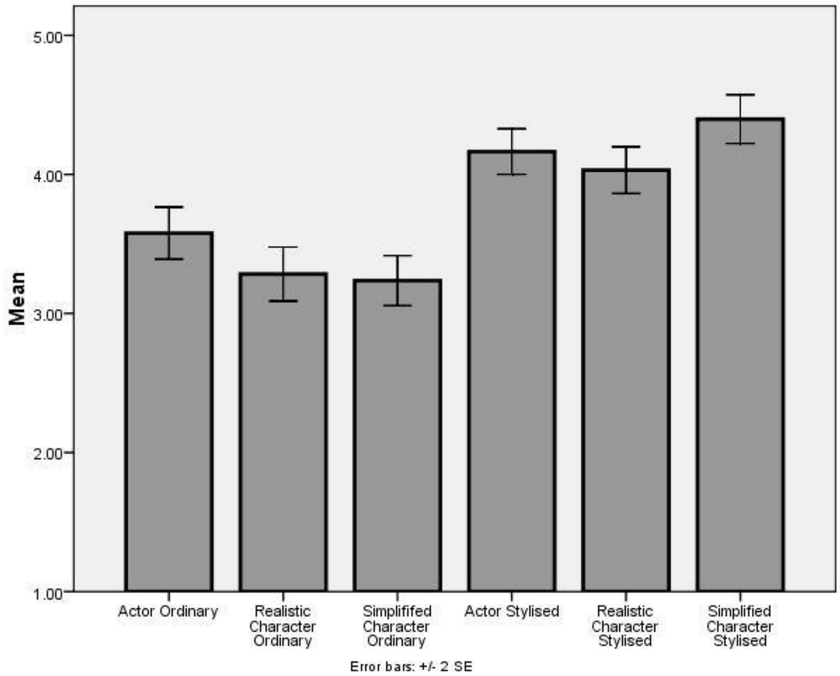


**Figure 5.9.** The overall *Strength* for the two styles for each character (Means + 2*Standard Error).

However, as suggested by the hypotheses (see Section 4.2), the results still showed significant interactions between *Character Type* and *Action Style* for *Strength* ($F(1,38)= 27.10$, $p<0.01$, Partial $\eta^2=0.41$). In other words, the effect of the *Character Type* on *Strength* differed depending on *Action Style.* For the Ordinary displays, there was a significant difference in *Strength* between the Actor, the Realistic Character and the Simplified Character ($F(1,38)=21.13$, $p<0.01$, partial $\eta^2=0.35$). Post-hoc analysis (Least Significant Difference) showed that *Strength* was significantly higher for the Actor than for the Realistic Character ($p<0.01$) and the Simplified Character ($p<0.01$). It was similar between the Realistic Character and the Simplified Character ($p<0.45$) (Figure 5.9). For the stylised displays, there was a significant difference in *Strength* between the Actor, the Realistic Character and the Simplified Character ($F(1,38)=14.59$, $p<0.01$, partial $\eta^2=0.27$) as well. However, post-hoc analysis (Least Significant Difference) showed that *Strength* was significantly higher for the Simplified Character than for the Actor ($p<0.01$) and the Realistic

Character ($p<0.01$). It was stronger for the Actor than for the Realistic Character ($p<0.05$) (Figure 5.9).

The trend to interaction that was found on *Believability* between *Character Type* and *Action Style* was reinvestigated as well.

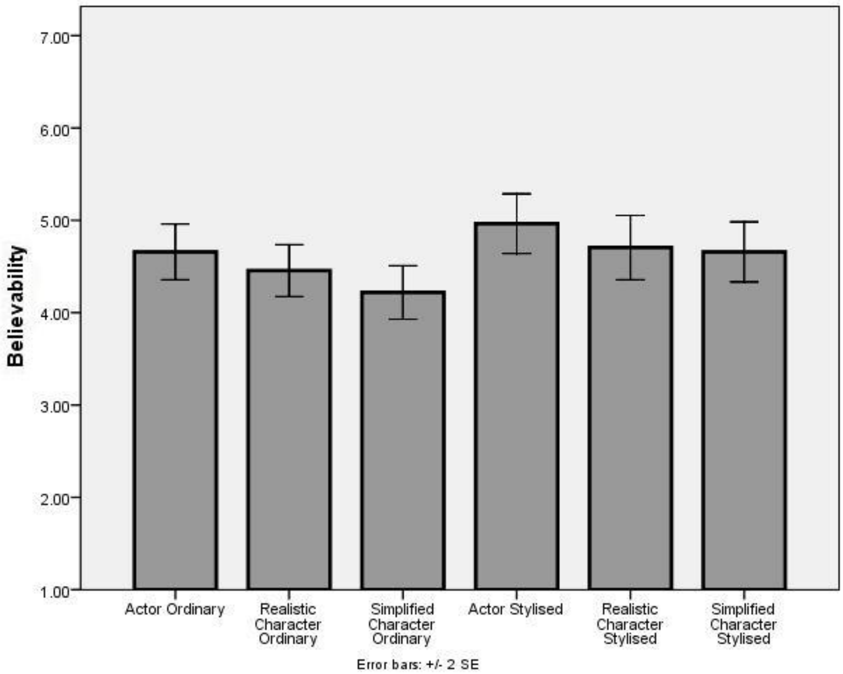


**Figure 5.10.** The overall *Believability* for the two styles for each character (Means + 2*Standard Error).

Although it did not reach significance, there was still a trend for *Believability* ($F(1,38)=2.68$, $p=0.08$, Partial $\eta^2=0.06$). For the Ordinary displays, there was a significant difference in *Believability* between the Actor, the Realistic Character and the Simplified Character ($F(1,38)=23.54$, $p<0.01$, partial $\eta^2=0.38$). Post-hoc analysis (Least Significant Difference) showed that *Believability* was significantly higher for the Actor than for the Realistic Character ($p<0.05$) and the Simplified Character ($p<0.01$). Moreover, it was higher for the Realistic Character than for the Simplified Character ($p<0.01$) (Figure 5.10). For the stylised displays, there was a significant difference in *Believability (B)* between the Actor, the Realistic Character and the Simplified Character ($F(1,38)=6.54$, $p<0.01$, partial $\eta^2=0.14$) as well. However, post-hoc analysis (Least Significant Difference) showed that *Believability* was significantly higher for the Actor than for Realistic Character ($p<0.01$) and the Simplified Character ($p<0.01$) but there was no difference in *Believability (B)* between the Realistic Character and the Simplified Character ($p=0.60$) (Figure 5.10).

Similarly, the interaction that was found on *Naturalness* between *Character Type* and *Action Style* was reinvestigated.

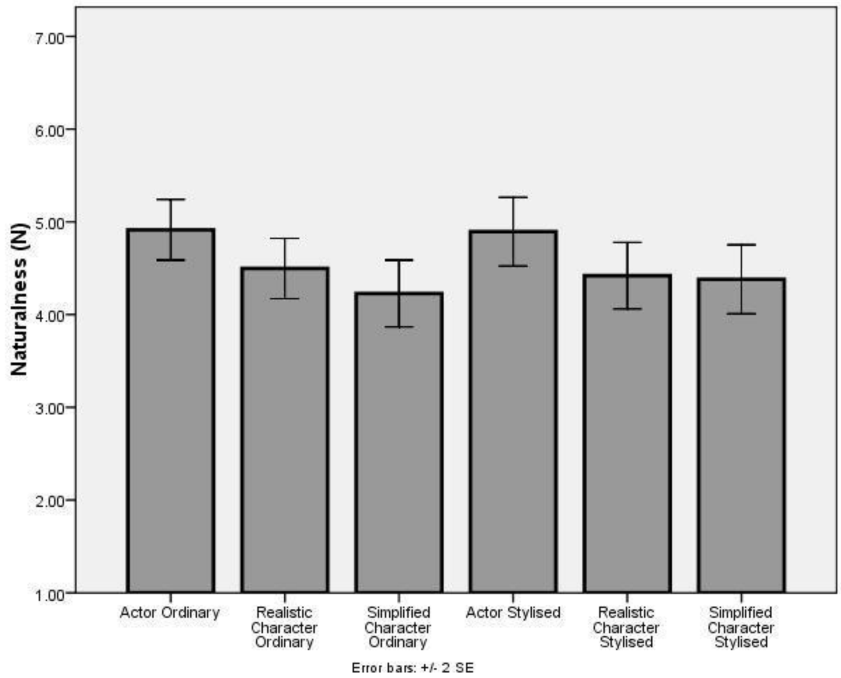


**Figure 5.11.** The overall *Naturalness* for the two styles for each character (Means + 2*Standard Error).

The results still showed an interaction for *Naturalness* ($F(1,38)= 3.46$, $p<0.05$, Partial $\eta^2=0.08$). For the Ordinary Displays, there was a significant difference in *Naturalness* between the Actor, the Realistic Character and the Simplified Character ($F(1,38)=15.48$, $p<0.01$, partial $\eta^2=0.28$). Post-hoc analysis (Least Significant Difference) showed that *Naturalness* was significantly higher for the Actor than for the Realistic Character ($p<0.01$) and the Simplified Character ($p<0.01$). Moreover, it was higher for the Realistic Character than for the Simplified Character ($p<0.01$) (Figure 5.11). For the stylised displays, there was a significant difference in *Naturalness* between the Actor, the Realistic Character and the Simplified Character ($F(1,38)=9.86$, $p<0.01$, partial $\eta^2=0.20$) as well. However, post-hoc analysis (Least Significant Difference) showed that *Naturalness* was significantly higher for the Actor than for Realistic Character ($p<0.01$) and the Simplified Character ($p<0.01$) but there was no difference in *Naturalness* between the Realistic Character and the Simplified Character ($p=0.69$) (Figure 5.11).

These combined results replace the original analysis (reported in Section 5.2). The detailed analysis of the effect of *Action Style* on *Correct Identification*, *Strength*, *Believability* and on

*Naturalness* for each emotion individually was also conducted on the combined data (i.e. both Frame rates together).

# 5.8. Effect of Action Style

## 5.8.1. Effect of *Action Style* on *Correct Identification*

| Emotions | Effect of *Action Style* on *Correct Identification* for the Actor | Effect of *Action Style* on *Correct Identification* for the Realistic Character | Effect of *Action Style* on *Correct Identification* for the Simplified Character |
|---|---|---|---|
| Anger | **F(1,38)=32.07, P<0.01, partial $\eta^2$ =0.46. O<S** | **F(1,38)=7.05, P<0.05, partial $\eta^2$ =0.16. O<S** | **F(1,38)=13.11, P<0.01, partial $\eta^2$ =0.26. O<S** |
| Disgust | **F(1,38)=41.76, P<0.01, partial $\eta^2$ =0.52. O>S** | **F(1,38)=25.69, P<0.01, partial $\eta^2$ =0.40. O>S** | **F(1,38)=57.60, P<0.01, partial $\eta^2$ =0.60. O>S** |
| Fear | F(1,38)=0.30, P=0.58, partial $\eta^2$ =0.01. | **F(1,38)=8.55, P<0.01, partial $\eta^2$ =0.18. O<S** | **F(1,38)=11.23, P<0.01, partial $\eta^2$ =0.23. O<S** |
| Sadness | F(1,38)=2.99, P=0.09, partial $\eta^2$ =0.07 | **F(1,38)=10.10, P<0.01, partial $\eta^2$ =0.21. O<S** | F(1,38)=3.74, P=0.06, partial $\eta^2$ =0.09. |
| Shame | **F(1,38)=11.40, P<0.01, partial $\eta^2$ =0.23. O<S** | **F(1,38)=12.09, P<0.01, partial $\eta^2$ =0.24. O<S** | **F(1,38)=19.14, P<0.01, partial $\eta^2$ =0.33. O<S** |
| Happiness | **F(1,38)=39.09, P<0.01, partial $\eta^2$ =0.51. O<S** | **F(1,38)=27.79, P<0.01, partial $\eta^2$ =0.42. O<S** | **F(1,38)=10.66, P<0.01, partial $\eta^2$ =0.22. O<S** |
| Excitement | F(1,38)=2.43, P=0.13, partial $\eta^2$ =0.06 | F(1,38)=0.90, P=0.34, partial $\eta^2$ =0.02 | **F(1,38)=7.12, P<0.01. partial $\eta^2$ =0.16. O<S** |
| Pride | F(1,38)=1.84, P=0.18, partial $\eta^2$ =0.05. | **F(1,38)=9.58, P<0.01, partial $\eta^2$ =0.20. O<S** | **F(1,38)=15.69, P<0.01, partial $\eta^2$ =0.51. O<S** |
| Relief | F(1,38)=2.68, P=0.11, partial $\eta^2$ =0.06. | F(1,38)=0.39, P=0.53, partial $\eta^2$ =0.01. | **F(1,38)=4.82, P<0.05, partial $\eta^2$ =0.11. R<O** |
| Surprise | **F(1,38)=27.32, P<0.01, partial $\eta^2$ =0.42. O<S** | **F(1,38)=34.08, P<0.01, partial $\eta^2$ =0.47. O<S** | **F(1,38)=22.77, P<0.01, partial $\eta^2$ =0.37. O<S** |

**Table 5.10.** Results of the repeated measures ANOVA carried out on *Correct Identification* between *Action Style* (O=Ordinary, S=Stylised) for each emotion displayed. The Actor, the Realistic Character and the Simplified character are separates.

To assess the effect of *Action Style* on *Correct Identification* for each Emotion Displayed, Repeated Measures ANOVAs were carried out (Table 5.10). The results show that for twenty-two of the thirty (10 emotions by 3 characters) Emotions Displayed, participants were significantly better in identifying the Stylised than the Ordinary movements for the Actor, the Realistic Character and for the Simplified Character. The only exception was the emotion "disgust" for which participants were better in correctly identifying the Ordinary display rather than the Stylised one for the three characters (Table 5.10). A post study of display materials, suggested this could be due to the body movements that the actor performed. In the Ordinary display, the actor performed as if he was keeping something far from his face whereas in the Stylised version he performed as if his body was covered by something. That could have made the Stylised version more difficult to interpret. Given this issue, it was decided to remove disgust from the list of emotions.

Thus, the results suggest that participants were better at interpreting the Stylised emotional body language than the Ordinary one.

### 5.8.2. Effect of *Action Style* on *Strength*

| Emotions | Effect of *Action Style* on *Strength* for the Actor | Effect of *Action Style* on *Strength* for the Realistic Character | Effect of *Action Style* on *Strength* for the Simplified Character |
|---|---|---|---|
| Anger | **$F(1,38)=73.89$, $P<0.01$, partial $\eta^2=0.66$. O<S** | **$F(1,38)=41.34$, $P<0.01$, partial $\eta^2=0.52$. O<S** | **$F(1,38)=31.03$, $P<0.01$, partial $\eta^2=0.45$. O<S** |
| Fear | **$F(1,38)=12.00$, $P<0.01$, partial $\eta^2=0.24$. O<S** | $F(1,38)=1.86$, $P=0.18$, partial $\eta^2=0.04$ | $F(1,38)=0.01$, $P=0.92$, partial $\eta^2=0.00$ |
| Sadness | $F(1,38)=2.75$, $P=0.11$, partial $\eta^2=0.07$ | **$F(1,38)=13.15$, $P<0.01$, partial $\eta^2=0.26$. O<S** | **$F(1,38)=34.25$, $P<0.01$, partial $\eta^2=0.47$. O<S** |
| Shame | **$F(1,38)=14.81$, $P<0.01$, partial $\eta^2=0.28$. O<S** | **$F(1,38)=15.92$, $P<0.01$, partial $\eta^2=0.29$. O<S** | **$F(1,38)=31.31$, $P<0.01$, partial $\eta^2=0.45$. O<S** |
| Happiness | **$F(1,38)=13.31$, $P<0.01$, partial $\eta^2=0.26$. O<S** | **$F(1,38)=34.74$, $P<0.01$, partial $\eta^2=0.48$. O<S** | **$F(1,38)=41.59$, $P<0.01$, partial $\eta^2=0.52$. O<S** |
| Excitement | **$F(1,38)=38.96$, $P<0.01$, partial $\eta^2=0.51$. O<S** | **$F(1,38)=13.95$, $P<0.01$, partial $\eta^2=0.27$. O<S** | **$F(1,38)=70.85$, $P<0.01$. partial $\eta^2=0.65$. O<S** |
| Pride | $F(1,38)=2.36$, $P=0.13$, partial $\eta^2=0.06$ | **$F(1,38)=9.65$, $P<0.01$, partial $\eta^2=0.20$. O<S** | **$F(1,38)=21.39$, $P<0.01$, partial $\eta^2=0.36$. O<S** |
| Relief | $F(1,38)=3.07$, $P=0.09$, partial $\eta^2=0.07$ | $F(1,38)=0.77$, $P=0.39$, partial $\eta^2=0.02$ | $F(1,38)=0.27$, $P=0.60$, partial $\eta^2=0.01$ |
| Surprise | **$F(1,38)=27.58$, $P<0.01$, partial $\eta^2=0.41$. O<S** | **$F(1,38)=43.90$, $P<0.01$, partial $\eta^2=0.54$. O<S** | **$F(1,38)=28.96$, $P<0.01$, partial $\eta^2=0.43$. O<S** |

**Table 5.11.** Results of the repeated measures ANOVA carried out on *Strength* between *Action Style* (O=Ordinary, S=Stylised) for each emotion displayed. The Actor, the Realistic Character and the Simplified Character are separates.

For each emotion, the effect of *Action Style* on *Strength* was assessed using repeated measures ANOVAs (Table 5.11). For twenty of the twenty-seven (9 emotions by 3 characters) Emotions Displayed, the Stylised emotion was interpreted as significantly stronger (*Strength)* than the Ordinary style for the Actor, the Realistic Character and the Simplified Character (Table 5.11). These results confirm the overall effect of *Action Style* on *Strength* that was found (see Section 5.4).

To summarise, although it was not the case for all the displays, participants found that the Stylised displays were stronger (*Strength*) than the Ordinary ones. An interaction between the *Action Style* and the *Character Type* was found (see Section 5.7). For the ordinary displays, participants found the actor more emotional than the realistic character, which was found more emotional than the simplified character. For the stylised displays, the simplified character was found more emotional than the Actor which was found more emotional than the Realistic Character (see Section 5.7).

### 5.8.3. Effect of *Action Style* on *Believability*

| Emotions | Effect of *Action Style* on *Believability* for the Actor | Effect of *Action Style* on *Believability* for the Realistic Character | Effect of *Action Style* on *Believability* for the Simplified Character |
|---|---|---|---|
| Anger | **$F(1,38)=15.19$, $P<0.01$, partial $\eta^2=0.28$. O<S** | $F(1,38)=2.04$, $P=0.16$, partial $\eta^2=0.05$ | **$F(1,38)=21.14$, $P<0.01$, partial $\eta^2=0.35$. O<S** |
| Fear | $F(1,38)=0.21$, $P=0.65$, partial $\eta^2=0.01$ | $F(1,38)=1.23$, $P=0.27$, partial $\eta^2=0.03$ | $F(1,38)=3.58$, $P=0.07$, partial $\eta^2=0.08$ |
| Sadness | $F(1,38)=0.29$, $P=0.59$, partial $\eta^2=0.01$ | $F(1,38)=3.12$, $P=0.08$, partial $\eta^2=0.07$ | **$F(1,38)=7.08$, $P<0.05$, partial $\eta^2=0.15$. O<S** |
| Shame | $F(1,38)=2.85$, $P=0.10$, partial $\eta^2=0.07$ | $F(1,38)=2.40$, $P=0.13$, partial $\eta^2=0.06$ | **$F(1,38)=11.69$, $P<0.01$, partial $\eta^2=0.23$. O<S** |
| Happiness | $F=0.13$, $P=0.72$, partial $\eta^2=0.00$ | $F(1,38)=0.40$, $P=0.53$, partial $\eta^2=0.01$ | $F(1,38)=2.28$, $P=0.14$, partial $\eta^2=0.06$ |
| Excitement | **$F(1,38)=5.87$, $P<0.05$, partial $\eta^2=0.13$. O<S** | $F(1,38)=1.69$, $P=0.20$, partial $\eta^2=0.04$ | **$F(1,38)=6.27$, $P<0.05$. partial $\eta^2=0.14$. O<S** |
| Pride | $F(1,38)=0.05$, $P=0.82$, partial $\eta^2=0.00$ | **$F(1,38)=6.16$, $P<0.05$, partial $\eta^2=0.14$. O<S** | **$F(1,38)=7.78$, $P<0.01$, partial $\eta^2=0.17$. O<S** |
| Relief | $F(1,38)=0.15$, $P=0.70$, partial $\eta^2=0.00$ | $F(1,38)=1.20$, $P=0.28$, partial $\eta^2=0.03$ | $F(1,38)=0.00$, $P=0.94$, partial $\eta^2=0.00$ |
| Surprise | **$F(1,38)=15.60$, $P<0.01$, partial $\eta^2=0.29$. O<S** | **$F(1,38)=9.76$, $P<0.01$, partial $\eta^2=0.20$. O<S** | **$F(1,38)=7.68$, $P<0.01$, partial $\eta^2=0.16$. O<S** |

**Table 5.12.** Results of the repeated measures ANOVA carried out on *Believability* between *Action Style* (O=Ordinary, S=Stylised). For the Actor, the Realistic Character and the Simplified character separately.

For each emotion, the effect of *Action Style* on *Believability* was assessed using repeated measures ANOVAs (Table 5.12). The pattern differs depending on the *Character Type*. For the Actor and the Realistic Character, there was only two to three of the nine Emotions Displayed for which an effect was found (Table 5.12). However, for the Realistic Character, the Stylised displays were always either found more *Believable* (six of the nine displays) or similar to the Ordinary display (Table 5.12). These results suggest that the overall effect of *Action Style* on *Believability* that was found (reported in Section 5.4) is due to a few exceptions for the Actor and the Realistic Character whereas for the Simplified Character it is the general pattern.

The detailed analysis suggests that, overall, *Action Style* did not affect the *Believability* of the Actor or the Realistic Character. However, for the Simplified Character, participants found that the Stylised displays were more *Believable* than the Ordinary ones. This is consistent with the interaction between *Action Style* and the *Character Type* that was found (see Section 5.7). For the Ordinary displays, participants found the actor more *Believable* than the Realistic Character, which was found more *Believable* than the Simplified Character. For the Stylised displays, the Actor was still was found more *Believable* than the Realistic Character and the Simplified Character. However, the Realistic Character and the Simplified Character were perceived as equally *Believable*.

## 5.8.4. Effect of Action Style on Naturalness

| Emotions | Effect of *Action Style* on *Naturalness* for the Actor | Effect of *Action Style* on *Naturalness* for the Realistic Character | Effect of *Action Style* on *Naturalness* for the Simplified Character |
|---|---|---|---|
| Anger | $F(1,38)=1.15$, $P=0.29$, partial $\eta^2=0.03$ | $F(1,38)=0.02$, $P=0.90$, partial $\eta^2=0.00$ | $F(1,38)=2.79$, $P=0.10$, partial $\eta^2=0.07$ |
| Fear | $F(1,38)=0.26$, $P=0.61$, partial $\eta^2=0.01$ | **$F(1,38)=7.84$, $P<0.01$, partial $\eta^2=0.17$. O>S** | **$F(1,38)=5.13$, $P<0.05$, partial $\eta^2=0.12$. O>S** |
| Sadness | $F(1,38)=1.90$, $P=0.18$, partial $\eta^2=0.05$ | $F(1,38)=0.42$, $P=0.52$, partial $\eta^2=0.01$ | $F(1,38)=1.57$, $P=0.22$, partial $\eta^2=0.04$ |
| Shame | $F(1,38)=0.12$, $P=0.73$, partial $\eta^2=0.01$ | $F(1,38)=0.94$, $P=0.34$, partial $\eta^2=0.02$ | $F(1,38)=0.54$, $P=0.47$, partial $\eta^2=0.01$ |
| Happiness | $F(1,38)=0.00$, $P=1.00$, partial $\eta^2=0.00$ | $F(1,38)=0.01$, $P=0.92$, partial $\eta^2=0.00$ | $F(1,38)=3.83$, $P=0.06$, partial $\eta^2=0.09$ |
| Excitement | $F(1,38)=2.56$, $P=0.12$, partial $\eta^2=0.06$ | $F(1,38)=0.44$, $P=0.51$, partial $\eta^2=0.01$ | $F(1,38)=1.86$, $P=0.18$, partial $\eta^2=0.04$ |
| Pride | $F(1,38)=2.05$, $P=0.16$, partial $\eta^2=0.05$ | $F(1,38)=0.44$, $P=0.51$, partial $\eta^2=0.01$ | $F(1,38)=1.57$, $P=0.22$, partial $\eta^2=0.04$ |
| Relief | **$F(1,38)=6.94$, $P<0.05$, partial $\eta^2=0.15$. O>S** | $F(1,38)=2.30$, $P=0.14$, partial $\eta^2=0.06$ | $F(1,38)=0.01$, $P=0.92$, partial $\eta^2=0.00$ |
| Surprise | **$F(1,38)=6.63$, $P<0.05$ partial $\eta^2=0.14$. O<S** | $F(1,38)=1.82$, $P=0.18$, partial $\eta^2=0.04$ | $F(1,38)=0.07$, $P=0.80$. partial $\eta^2=0.00$ |

**Table 5.13.** Results of the repeated measures ANOVA carried out on *Naturalness* between *Action Style* (O=Ordinary, S=Stylised). For the Actor, the Realistic Character and the Simplified character separately.

For each emotion, repeated ANOVAs were carried on the data. For the majority of emotions *Action Style* had no significant effect on *Naturalness* (Table 5.13). These results confirm the overall lack of effect of *Action Style* on *Naturalness* that was found (see Section 5.4).

Overall, participants found that the Stylised displays were equally *Natural* as the Ordinary displays. However, an interaction between the *Action Style* and the *Character Type* was found previously (see Section 5.7). For the Ordinary displays, participants found the Actor more *Natural* than the Realistic Character, which was found more *Natural* than the Simplified Character. For the Stylised displays, the Actor was still found more *Natural* than the Realistic Character and the Simplified Character. However, the Realistic Character and the Simplified Character were perceived as similar regarding *Naturalness* (see Section 5.7).

## 5.8.5. Summary of the effects of *Action Style*

Participants were better at interpreting the Stylised emotional displays than the Ordinary ones for the Actor, the Realistic Character and the Simplified Character.

Overall, the Stylised displays were perceived as emotionally stronger than the Ordinary ones. Moreover, an interaction was found between *Character Type* and *Action Style*. For the Ordinary displays, the Actor was perceived as emotionally stronger (*Strength*) than the

Realistic Character, which was found emotionally stronger *(Strength)* than the Simplified Character. For the Stylised Displays, the Simplified Character was found emotionally stronger (*Strength*) than the Actor, which was found emotionally stronger *(Strength)* than the Realistic Character.

*Action Style* had an effect on *Believability*, the Stylised displays were found more believable than the Ordinary ones for the Simplified Character. However, the detailed analysis showed that this was not generally the case for the Actor and the Realistic Character.

Overall, *Action Style* had no effect on *Naturalness*. There was an interaction between *Action Style* and *Character Type*. For the Ordinary displays, the Actor was found more *Believable* and more *Natural* than the Realistic Character, which was found more believable and more natural than the Simplified Character. However, for the Stylised displays the Actor was still found more *Believable* and more *Natural* than the two animated characters but the Realistic Character was found as *Believable* and *Natural* as the Simplified Character.

## 5.9. Analysis of Individual Differences

The relationship between *Emotional Intelligence* and *Correct Identification*, was computed using a Pearson correlation. There was no correlation between *Emotional Intelligence* and *Correct Identification* for the Actor ($r(38)=0.18$, $p=0.28$), or for the Realistic Character ($r(38)=0.15$, $p=0.35$), or for the Simplified Character ($r(38)=0.15$, $p=0.55$).

The relationship between *Emotional Intelligence* and *Strength*, was computed using a Pearson correlation. There was a positive correlation between *Emotional Intelligence* and *Strength* for the Actor ($r(38)=0.40$, $p<0.05$), for the Realistic Character ($r(38)=0.51$, $p<0.01$), but not for the Simplified Character ($r(38)=0.23$, $p=0.15$).

Similarly, correlations were computed for *Emotional Intelligence* and *Believability*. There was no relationship between *Emotional Intelligence* and the *Believability* of the Actor ($r(38)=0.23$, $p=0.16$) nor with the Simplified Character ($r(38)=0.21$, $p=0.19$). However, *Emotional Intelligence* was positively correlated with the *Believability* of the Realistic Character ($r(38)=0.33$, $p<0.05$). The results are similar with the *Naturalness*, as there was no correlation between *Emotional Intelligence* and the *Naturalness* of the Actor ($r(38)=0.20$, $p=0.20$) nor with the *Naturalness* of the Simplified Character ($r(38)=0.24$, $p=0.15$).

However, *Emotional Intelligence* was positively correlated with the *Naturalness* of the Realistic Character ($r(38)=0.32$, $p<0.05$).

Pearson's correlations also showed a positive correlation between *Experience in Video Games* and *Correct Identification* for the Actor ($r(38)=0.55$, $p<0.01$) as well as the *Correct Identification* of the Simplified Character ($r(38)=0.35$, $p<0.05$) but not with the Realistic Character ($r(38)=0.18$, $p=0.26$). However, *Familiarity with Animated Characters* was not correlated with *Correct Identification* for the Actor ($r(38)=0.17$, $p=0.29$), the Simplified Character ($r(38)=0.15$, $p=0.37$) nor with the Realistic Character ($r(38)=0.27$, $p=0.09$).

There was a positive correlation between *Experience in Video Games* and *Strength* for the Actor ($r(38)=0.33$, $p<0.05$) but not for the Realistic Character ($r(38)=0.29$, $p=0.06$) nor for the Simplified Character ($r(38)=0.28$, $p<0.08$). However, *Familiarity with Animated Characters* was not correlated with *Strength (S)* for the Actor ($r(38)=0.14$, $p=0.37$), for the Realistic Character ($r(38)=0.06$, $p=0.72$) nor for the Simplified Character ($r(38)=0.24$, $p=0.88$).

Interestingly, in the post hoc questionnaire, 9 out of 40 participants reported that during the study the Simplified Character and the Realistic Character were actors and 15 out of 40 participants reported that the Simplified Character was animated and that the Realistic Character was an actor. Thus, only 16 out of 40 participants correctly identified the different character presentation conditions.

# Chapter 6 - A Comparative Study of Emotional Body Language Displayed by Humans and Animated Characters: Discussion of the results

## 6.1. Introduction

The study was investigating four questions (see Section 4.2.4) that are discussed in this chapter.

(Q1) Does the *Character Type* affect the *Correct Identification*, *Strength*, *Believability* and *Naturalness* of the emotional body language displayed?

(Q2) Does the *Action Style* affect the *Correct Identification*, *Strength*, *Believability* and *Naturalness* of the emotional body language displayed?

(Q3) Does the *Frame Rate* affect the *Correct Identification*, *Strength*, *Believability* and *Naturalness* of the emotional body language displayed?

(Q4) Are personal differences including *Emotional Intelligence*, *Experience in Playing Video Game* or *Familiarity with Animated Character* related to *Correct Identification*, *Strength, Believability* and *Naturalness*?

## 6.2. Character Type does not affect the Correct Identification of the emotional body language displayed.

The questionnaire used to record participants' *Correct Identification* provided them with twenty-two different emotional labels making it hard to choose an emotion by elimination. Participants were better at correctly identifying the Emotions Displayed than expected by chance (Table 5.1) even though faces and hands were hidden (Figure 4.1). It can therefore be concluded that it is possible to correctly interpret the emotions that were displayed through watching body language only.

Considering the emotional content for each condition, there was no significant difference in *Correct Identification* of the emotional body language displayed by the Actor, the Realistic Character and the Simplified Character for most of the Emotions Displayed (see Section 5.5.1). However, for seven displays out of twenty-one the *Character Type* had an effect on *Correct Identification* (Table 5.6). Nevertheless, the overall similarity of identification for the different characters was confirmed by the clustering of the Emotions Displayed in the psychological space (Figure 5.4) resulting from the multi dimensional scaling (see Section 5.5.1). In Figure 5.4, displays that were similarly labelled are 'close' to each other. The configuration suggests that *Correct Identification* was consistent across different *Character Type* conditions and style displayed. The psychological space also confirms that the few differences found in *Correct Identification* between the different characters were exceptions. "Fear Ordinary", for example, was correctly identified more often when displayed by the Actor (Table 5.6) than by the Realistic Character or the Simplified Character. However, the psychological space (Figure 5.4) indicates that the emotions with which it was confused when displayed by the animated characters were close across different *Character Type*. The psychological space shows similar results for the display of "Disgust Stylised", "Relief Ordinary", "Happiness Ordinary", "Excitement Stylised" that varied in *Correct Identification* across *Character Type* but were still close in the psychological space (Figure 5.4). Moreover, Figure 5.4 highlights strong confusion between the displays of "Shame" and "Sadness" that were often correctly identified but also often confused with "Guilt" and "Disappointment". This overall similarity is consistent with the results of the previous study (see Section 3.5.3) as well as with (McDonnell, Jorg, McHugh, Newell, & O'Sullivan, 2008) who found that overall participants are not affected by the characters' 'look' (i.e. *Character Type*) when identifying emotions. Considering the psychological space and the similar results that have been found in these two other studies, it is becoming unlikely that the lack of difference in recognition rate for most of the displays is due to weak statistical power or to the difficulty to correctly identify emotions watching body language only (floor effect). Instead, *Correct Identification* of the emotional body language displayed was similar across the three *Characters Types*. It can be concluded that neither the simplification of the level of physical realism or the nature of the character (Human vs. Artificial) have degraded the emotional 'channel' of information for most of the displays.

Although, participants were not broadly affected by the *Character Type* for identifying the emotions (*Correct Identification*), there were seven out of twenty one Emotions Displayed for which significant differences were found (Table 5.6). For four of them there were significant differences between the Simplified Character and the two other characters (Table 5.6). Thus, it can be concluded that, for these four displays the change in the level of physical realism had an effect on *Correct Identification*. Moreover, even though the actor tried to display happiness ordinarily, the emotional body language displayed could have been perceived as stylised as the actor was laughing and clapping. Therefore, these differences between the characters could have resulted from not following one of the key principles of animation. In traditional animation, different characters should have different personalities and should express emotions differently (Thomas & Johnston, 1995). In other words, the emotional body language displayed by a character should be consistent with its personality and with the way it looks etc (Appendix A). When watching "Happiness Ordinary" and "Excitement Stylised" participants could have found the emotional body language displayed by the Actor and the Realistic Character inconsistent with the way these two characters look. This inconsistencies could have resulted in the emotions displayed being unconvincing and more difficult to interpret, whereas, the same body language displayed by the Simplified Character could appear consistent with its "look" and hence resulted in a significantly higher number of *Correct Identification* (Table 5.6).

However, the same animation principle does not explain why participants were better in identifying three of the emotional displays when presented by the Actor (Table 5.6): The differences between the Actor and the Realistic Character would not be expected as more than half of the participants did not notice that it was an animation (Section 5.9). Furthermore, for these displays, there was no difference between the Realistic Character and the Simplified Character. This lack of difference suggests that the level of physical realism did not have an effect on the number of *Correct Identification*. Taken together, these results suggest that the difference in recognition rate for these displays were resulting from the missing micro gestures and secondary animation rather than the change of *Character Type*. Moreover, this is consistent with another animation principle which emphasizes the importance of secondary animations (Thomas & Johnston, 1995) (see Appendix A).

## 6.3. Character Type affects the Strength of the emotional body language displayed.

Overall, *Character Type* had an effect on *Strength*. The emotions were perceived as stronger (*Strength*) when displayed by the Actor than when displayed by the animated characters (see

Section 5.4). However, for a large majority (fifteen out of twenty one) of Emotions Displayed, there was no difference in emotional *Strength* between the three *Character Types* (see Section 5.5.2). This is consistent with the finding that overall *Correct Identification* was not affected and rather suggests that emotional interpretations (*Correct Identification* and *Strength*) were similar across the different characters for most of the displays.

However, there were six (out of twenty one) Emotions Displayed for which significant differences in emotional *Strength* between the actor and the two animated characters were found (Table 5.7). These differences could be due to the change in the level of physical realism manipulated by changing the *Character Type*. This would explain why significant differences were found between the Realistic Character and the Simplified Character when displaying "Shame Ordinary" and "Pride Ordinary" (Table 5.7). The same motion capture data were used for both characters; therefore the only difference between the two animations was the *Character Type*. Thus it can be concluded that at least for these two displays the change in the level of realism had an effect on the perceived *Strength*. These differences between the Realistic Character and the Simplified Character are consistent with traditional principles of animation. In traditional animation, each character should express emotions in its own way (Thomas & Johnston, 1995). It is therefore possible that when watching the Simplified Character expressing "Shame Ordinary" and "Pride Ordinary" participants found the emotion weaker and less convincing as the body language used may not work well for the physical aspect of this character.

The significant differences that were found in the four others emotional displays seem dissimilar. There was no difference in emotional *Strength* between the Realistic Character and the Simplified Character (Table 5.7). If the differences in emotional *Strength* were due to the level of physical realism then a similar difference between the two animated models would be expected. On the other hand, the loss of secondary cues and micro gestures, introduced by the motion capture technology would explain why, for some displays, the emotional body language displayed by the Actor was found to be emotionally stronger than the Realistic and Simplified Characters and also why there was no difference between the two animated models. For example, in displaying emotions such as relief and sadness, the actor used visible sighs along with other movements. These sighs are either not present or heavily diminished when the emotions were displayed by the Realistic Character and the Simplified Character. Since breathing might be an important cue to interpret the strength of an emotion

(Dantzer, 2005), it could explain why these emotions were still correctly identified but were perceived as less emotional when displayed by the animated characters. Therefore, the results suggest that for these four displays, the missing micro gestures and secondary animations did have an effect on the perceived emotional *Strength*.

## 6.4. Character Type affects the Believability and Naturalness of the emotional body language displayed.

Overall, the Actor was found to be more *Believable* and more *Natural* than the Realistic Character which in turn was found to be more *Believable* and more *Natural* than the Simplified Character when displaying emotions ordinarily (see results reported in Section 5.5.3 and Section 5.5.4). Once again, these differences could be due either to different level of realism or to the loss of secondary cues introduced by the motion capture technology. The emotional body language displayed by the Realistic Character and by the Simplified Character was animated using the same motion capture data. Therefore the secondary animations and micro gestures that are missing are exactly the same and could not make any difference between them. Hence, the significant differences between the two animated models suggest that the level of physical realism has an effect on *Believability* and *Naturalness*.

## 6.5. Action Style affects the Correct Identification, Strength, Believability and Naturalness of the emotional body language displayed.

As expected, *Action Style* made a significant difference in *Correct Identification*. The stylised displays were easier to identify and thus participants were better in interpreting them. Overall, the effect is similar for the perceived *Strength*, for which the stylised displays were found more emotional for the three *Character Types*. For *Believability, Action Style* had an effect on the Emotions Displayed by the Simplified Character only, for which the Stylised displays were found more *Believable* than the Ordinarily ones.

However, an interaction between *Character Type* and *Action Style* was found. When displaying the emotions ordinarily, the Actor was found more emotional (*Strength*) and more believable (*Believability*) than the Realistic Character, which in turn was found more emotional and believable than the Simplified Character (see Section 5.7). This suggests that the body of the Realistic Character is more adapted than the body of the Simplified Character in order to display emotions ordinarily. However the results were different when displaying the stylised emotional body language. For these displays, the Simplified Character was found more emotional than the Actor which was found more emotional than the Realistic Character (see Section 5.7). This is consistent with the animation principle (Appendix A) which states that animated characters should move accordingly with the way they look (Thomas & Johnston, 1995). The lack of realism of the Simplified Character makes it look more 'cartoony' (Figure 4.2). Therefore the fact that it was found more emotional than the Actor and the Realistic Character when displaying stylised emotions but less emotional than the other two when displaying emotions ordinarily is consistent with the same animation principle. Moreover, when displaying stylised emotions, there was no difference in *Believability* for the Simplified Character and the Realistic Character which confirms that, overall, the Simplified Character is better than the Realistic Character for displaying stylised emotions. These findings strongly suggest that the Uncanny Valley results from a discrepancy between the way characters look and the way they move.

There was no effect of *Action Style* on *Naturalness* for a large majority of displays (Table 5.13). This result and the fact that *Character Type* had an effect suggest that perceived *Naturalness* is independent from the body language displayed. It shows that the important factor is the physical appearance of the character displaying the body language and not the actual body language itself.

## 6.6. Movements' Smoothness does not affect the Correct Identification, Believability and Naturalness of the emotional body language displayed.

The change in frame rate did not affect *Correct Identification*, *Strength*, *Believability* or *Naturalness*. Regarding *Naturalness*, the lack of difference is consistent with the findings that suggest that naturalness is independent from the body language itself. However, the lack of effect in the other dependent variables shows that the manipulation did not affect the 'channel' of emotional information (*Correct Identification* and *Strength*) nor the way the character was perceived.

This condition was introduced to reproduce the loss due to the Motion capture technology and is consistent with the findings that for most of the displays the ‘channel’ of emotional information was not affected by the micro gestures and secondary animation that were lost. Moreover, in both conditions participants were equally good in identifying a wide range of emotion including complex emotional states such as pride or shame.

## 6.7. Individual Differences are related to Correct Identification, Strength, Believability and Naturalness.

The positive correlation that was found between *Experience in Playing Video Games* and the *Believability* of realistic character may well be due to participants’ being used to interacting with realistic characters (e.g. physically accurate). If this is the case it would imply that as users’ get used to highly realistic characters, they find them more believable and that the feeling described by the Uncanny Valley may be linked to the novelty of the situation.

Moreover, as in Chapter 3, significant positive correlations were found between *Emotional Intelligence* and *Believability* of the Realistic Character. In other words, participants who scored high in *Emotional Intelligence* rated the Realistic Character as more believable. The realistic character was the one most likely to be confronted with the Uncanny Valley. Therefore, this could indicate that participants who score high to the questionnaire are less likely to be affected by the drop described by the Uncanny Valley while interacting with realistic characters.

These two correlations suggest some limitations for applying the Uncanny Valley model to animation. The Uncanny Valley models a two dimensional equation between human likeness (i.e. how human it looks) with how familiar (believable) it feels. Considering the position of the characters on the Uncanny Valley, the only possible conclusion is that, even though it was found less believable and less natural than the Actor, the Realistic Character did not ‘fall’ into the valley as it was still found more believable and more natural than the Simplified Character for the Ordinary displays. This does not necessarily imply that the model is inaccurate (further discussed in Section 7.3) as it could be that the Realistic Character is near the ‘top’ of the cliff or could be situated on the right of the valley (Figure 2.1). However, the perceived believability of an animated character varied depending on emotional intelligence and experience in video games (and possibly other individual differences) and hence a two dimensional equation representing the Uncanny Valley problem seems to be incomplete. In other words, the position of a specific character on the Uncanny Valley graphic (Figure 2.1) varies among different users. Thus, if an alternative model has to be proposed it will have to consider not only the physical realism of the character but also users' personal differences in order to predict whether an animated character will be perceived as believable by a specific audience.

## 6.8. Conclusion

Overall the identification of emotional body language was found to be unaffected by degrading the videos nor by changing the biological character to a computer generated one. The results suggest that the difference in identification that was found for the negative emotions in the first study was mostly due to the missing secondary animations. In the previous study, hands and fingers movements were not hidden. This resulted in the actor's fingers being in movements whereas the animated character's fingers did not move. This was not the case for this study as the actor's and animated characters' hands were completely hidden. This can explain the different results found in the two studies.

Nevertheless, some critical differences in interpretation were found at different level of realism. The Realistic Character was found to be more adapted for displaying emotions ordinarily whereas the Simplified Character was found to be better for displaying stylised emotions. These differences verify the traditional animation principle that highlights the importance of maintaining consistency between characters' movements and the way it looks (Appendix A). Regarding the training of medical students for interviewing (The area of interest for this project – Section 1.3), it suggests that realistic characters should be preferred, as they will have to display emotions in an ordinary and believable way. The problem is that they will be confronted with the loss in believability that was found between the biological character and the computer generated ones. Moreover, the differences that resulted from the motion capture technology are consistent with another traditional animation principle which highlights the importance of secondary animation. Therefore, if motion capture technology is used, the animation will have to be completed to include cues such as breathing that may be important for the emotional interpretation of the body language. Moreover, adding the missing secondary animation could diminish the differences that were found between the biological and the animated character. However, this will raise issues regarding the realism of the secondary and micro gestures that will have to be animated (rather than recorded) and which could provide trainee with cues that would not be present in a real life situation.

# CHAPTER 7 – CONCLUSIONS

## 7.1. Research Findings

The aim of this research programme was to assess the effect of realism on the perception of emotional body language displayed by animated characters (see Section 1.6.1 for further discussions). More specifically, the high level question addressed in this thesis was: ***Does behavioural and/or physical realism affect the perception of animated emotional body language?***

The main finding is that emotional body language displayed by animated characters can be accurately identified (see discussions in Section 3.6.1 and Section 6.2) and that the recognition rate of an emotion is not affected by the nature of the character (i.e. biological or computer generated) displaying the emotional body language (see Section 6.2). In other words, it is possible for animated characters to successfully express emotions using their bodies. However, it was found that both physical realism and behavioural realism (i.e. the way a character moves) affects the perceived emotional strength, believability and naturalness of the body language displayed (see Section 6.3 and Section 6.4). Differences were found between biological and computer generated characters. These differences suggest that behavioural realism and physical realism should be consistent (see Section 6.5). This has strong implications for the design of Virtual Environments for social task training.

## 7.2. Implications for Social Task Training

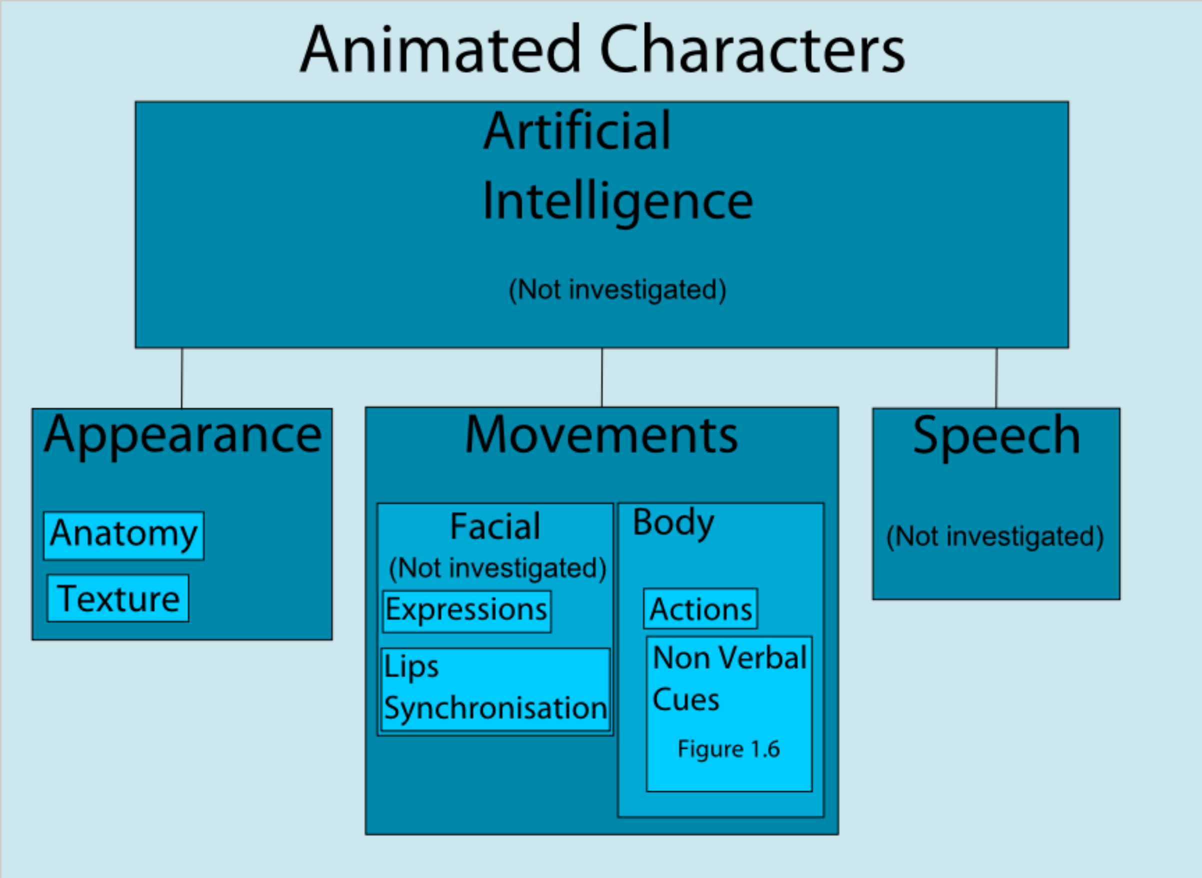


**Figure 7.1** Elements specific to the animated characters within a Virtual Environment.

Although differences were found in terms of emotional strength, believability and naturalness, the finding that realism of anatomy (Figure 7.1) does not affect the identification of an emotion initially supports the suitability of virtual characters for social task training (see Section 6.2). For instance, the results suggest that an animated character presenting emotional distress can be interpreted as such (providing it behaves as a human would do in such a context).

Moreover, the findings suggest that consistency between the body anatomy of a character and the body language it displays (Figure 7.1) is an essential component for believability. For social task training, the characters should behave realistically in order to avoid learning simplified clues. Therefore, the physical appearance should match the behavioural realism necessary for such training. In other words, for social task training, realistic characters should be used in order to maintain consistency between physical realism and the behavioural realism needed for the training.

Thus, the problem introduced by the Uncanny Valley cannot be avoided by using a physically unrealistic character behaving like a human. This raises technical issues both in Artificial Intelligence, which will eventually have to model highly complex states and in animation, which will have to model rich emotional body language in order to still maintain believability. However, even a realistic character, behaving consistently with the way it looks, may be subject to a loss of emotional strength, believability and naturalness as was found between the human and the realistic character (see Section 6.4).

It was not possible to draw any conclusion regarding whether these losses would undermine Social Presence (see Section 1.2) within a Virtual Environment for social task training. If this is the case, the results suggest that the problem could be solved by increasing the intensity of the emotion being displayed. In other words, to solve this reduction in strength etc, it could be possible to use an unrealistic character and increase the intensity of the emotions displayed (it was found that stylised animated display of emotions are perceived as more believable). However, stylised emotions would lead to unrealistic situations (in the social task training context) and thus to the learning of skills that may not be transferable to real life situations, i.e. inaccurate transfer of training (see Section 1.2).

Thus, since behavioural realism (i.e. not stylised emotions) is necessary for the training of social skills (e.g. diagnostic interview), the character's appearance will also have to be realistic as it was found that they should be consistent (see Section 6.5). Thus, given the loss in believability that was found and the necessity to display realistic behaviours, future research should investigate whether it is possible for such characters to trigger Social Presence (see Section 7.6).

## 7.3. Implications for the Uncanny Valley

The results suggest that the Uncanny Valley results from a mismatch between the way a character looks (i.e. its anatomy) and the emotional body language it displays (see Section 6.5). This is consistent with the theory that as characters look or behave realistically, viewers raise their level of expectation, which induces a drop in believability if the physical appearance and behaviour does not match (see Section 2.5). This is similar to the suggestion made by Mori (1970) with respect to robots.

It should be noted that the shape of the Uncanny Valley makes it a challenge to empirically investigate. The main study (reported in Chapter 4, 5, 6) used three levels of physical realism and two levels of behavioural realism. However, it is impossible to establish with certainty where the position of the characters are.

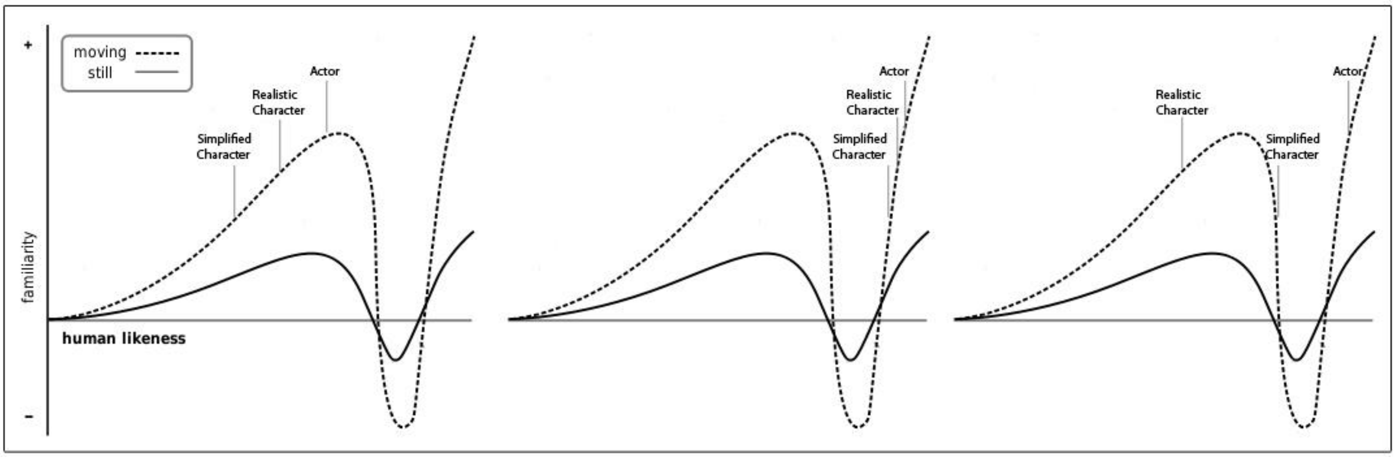


**Figure 7.2** Three possible positions of the characters tested on the Uncanny Valley

For example, the same patterns could occur if the three characters were on the left of the valley (Figure 7.2 - Left), on the right (Figure 7.2 - Middle) or with one character to the left, one character near the fall and one to the right (Figure 7.2 - Right). This suggests that using a wider range of physical realism levels, as done in existing studies (Section 2.3.2), would be an appropriate approach. However, the interactions that were found between behavioural and physical realism suggests that such an approach fails to completely capture the nature of the problem because they do not consider dynamic displays. The results in this thesis suggest that believability is given by a combination of physical realism, behavioural realism as well as individual participant differences.

Moreover, there is also a major issue when it comes to applying this theory to animation. Mori's aim was to describe a problem that occurs as robots get more realistic. It is not evident that researchers working in the robotic fields and in the animation field are actually working on the same issue. The Uncanny Valley is typically described in animation as a drop in believability. In robotics, however, it is described as a repulsive feeling. It is of course possible that the repulsive feeling is indeed a stronger version of the drop in believability, which could be due to the fact that a robot is "present" in the real world. However, this is not evident and the problem could have different causes and shape.

Hence, exploring the causes and shape of the Uncanny Valley theory is an important issue. Future work will benefit from the results of this thesis as the results show that the Uncanny Valley model is incomplete. It was found that individual differences in participants' Emotional Intelligence are related to the perceived believability of a character (see Section 3.5.7 and Section 5.9). It would be interesting to investigate if other personal differences are related as this could lead to a better understanding of the causes of the Uncanny Valley. Hence, the results suggest that in order to fully understand the phenomenon, it will be necessary to investigate the effect of physical and behavioural realism as well as individual participant differences and how they interact with the perception of realistic animated characters (see Section 7.6).

## 7.4. Implications for Psychology

It was previously established that body language can be accurately identified independently from facial and vocal expressions. However, most of the studies used a limited set of emotions (Atkinson, et al., 2004; den Stock, et al., 2007; Kleinsmith, et al., 2006). Moreover, these studies used a forced choice questionnaire between a very limited set of answers. This could have resulted in participants discriminating between the different emotions rather than identifying them (Winters, 2005). Moreover, due to the limited set of emotions used, the results do not necessarily imply that it is possible to interpret a wide range of emotions. However, using the emotional wheel as the questionnaire in both studies (Chapter 3 and Chapter 4) offered a wider choice of emotions, confirming that participants were able to recognize the different emotions presented. Moreover, the material used for the studies reported in this thesis included a wider set of emotions and included social emotion such as shame and pride. Participants were able to correctly interpret them when displayed through body language only. Hence, the recognition rates obtained (see results reported in Section 3.5.3 and Section 5.1) show that the expressivity of body language can be extended beyond the basic emotions. Participants were able to identify a set of ten emotions watching body language only. However, these results were established using acted emotional body language. It would be interesting to test whether it is possible to obtain similar results using real life displays.

Since the nature (computer generated or biological) of a character does not affect the recognition of an emotion, animated characters can be used effectively in psychology studies. For instance, Motion Capture technology could be used to systematically study the effect of context on the interpretation of emotional body language. It was used throughout this thesis to investigate the effect of changing the physical appearance of the character displaying it. Similarly, since the nature (computer generated or biological) of a character does not affect the recognition of an emotion, the effect of changing the context could be investigated by displaying the same emotional body language in different environments. This could lead to a better understanding of the perception of emotional body language and how and when it is displayed. Ultimately, this knowledge would benefit Virtual Environment for social task training as it is necessary to portray characters within the context of a situation and their behaviours should be consistent with the context in which it is displayed.

## 7.5. Confirmation of Results with Robotics

The theory of the Uncanny Valley is controversial, yet the results from this thesis are internally consistent, and consistent with Mori's (1970) model. However further confirmation of their validity has come from several studies conducted subsequently in the related domain of Robotics (as per Mori).

The studies reported in Chapters 3-6 confirmed that it is possible to accurately identify emotions displayed by body language alone, and found that the physical appearance of the body does not affect the identification of an emotion. Moreover, regarding movements, it was found that stylized emotional body language, comparable to the one that can be seen in animation, is perceived as more believable than ordinary displays and that a character should move consistently with the way it looks. Thus, it suggested that even simplified humanoid robots can use body language to express emotions.

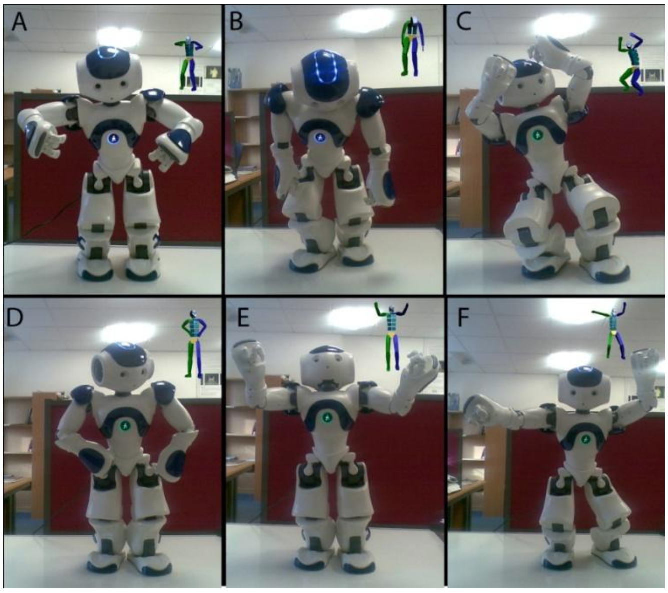


**Figure 7.3** The six key poses (A: Anger, B: Sadness, C: Fear, D: Pride, E: Happiness, F: Excitement)

This presents an opportunity for many humanoid robots, for example Nao (Figure 7.3), as they do not have the ability to display facial expressions. Hence, their bodies are the only medium available to express emotions. The results from this thesis suggested that emotional body language is an appropriate medium for expressing emotions, even with simplified bodies.

These results have since been confirmed for robotics within the FEELIX GROWING project, which was looking at endowing robots with emotions. Two experiments using the material created for the study reported in Chapter 4 extended the results from this Ph.D by finding that body language can be accurately perceived when displayed by a robot (Beck, Cañamero, & Bard, 2010; Beck, Hiolle, Mazel, & Cañamero, 2010). Figure 7.3 shows the key poses used to investigate how emotional body language displayed by a humanoid robot is interpreted. Each panel of Figure 7.3 show the key pose in Nao (Beck, Cañamero, et al., 2010) and the original key pose, taken from the Motion Capture data used in Chapter 4, which served as a model (top right of the image). The first goal of the study was to test the expressivity of the key poses displayed by the robot. The results show that participants were far better than chance level at interpreting the different key poses taken by the robot. The results also showed that the interpretation of the body language was consistent with the actor's intentions while performing. Thus extending the results presented in this thesis. This strand of research is now extended within the FP7 ALIZ-E European project.

# 7.6. Recommendation for Future Work

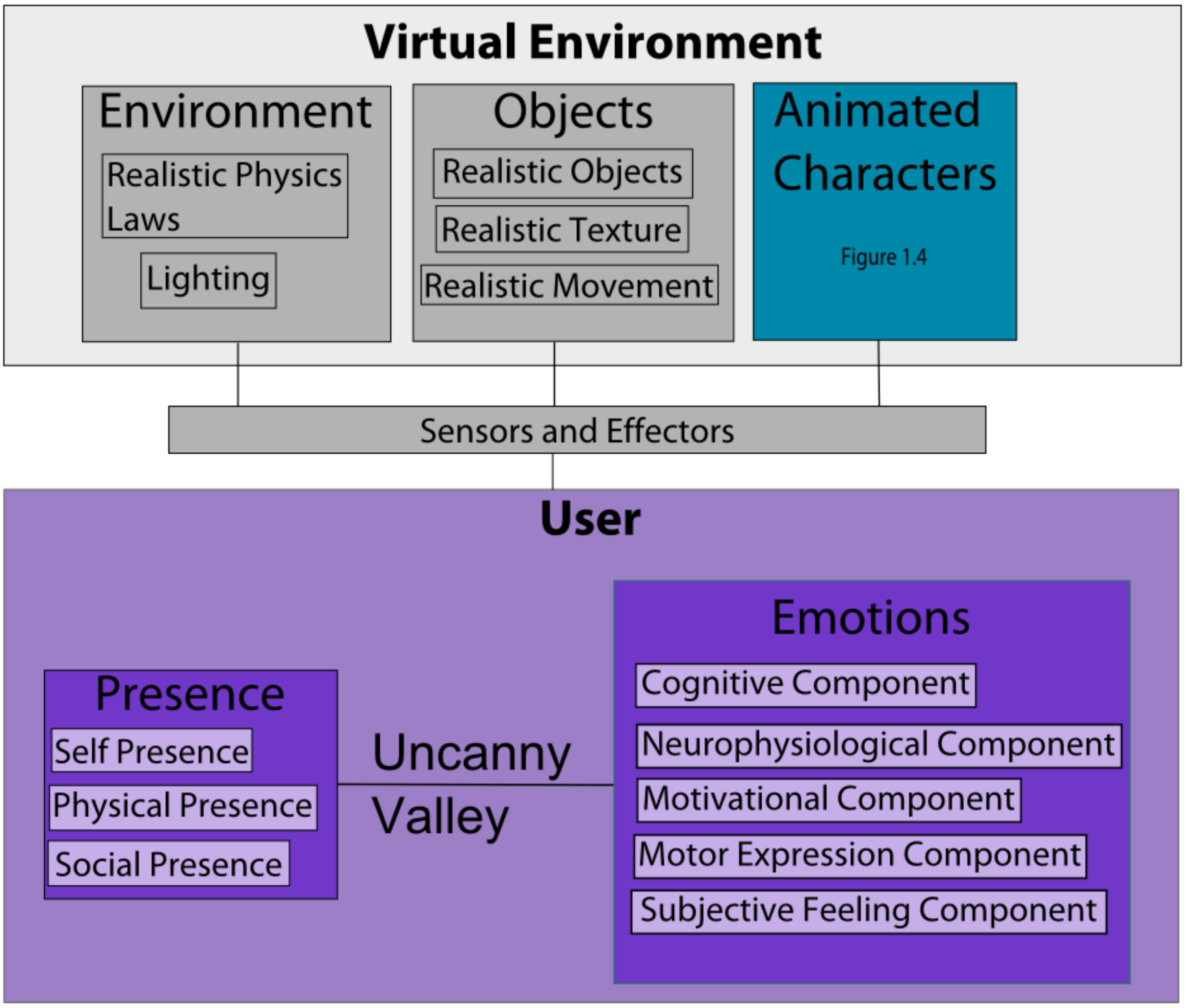


**Figure 7.4** Elements required for accurate transfer of skills lead to the Uncanny Valley.

## 7.6.1. Effect of the loss of believability on Social Presence

To be effective for social task training, Virtual Environments need to induce Presence (see Section 1.2 for definition). However, the loss of believability that was found could result in the animated characters not inducing Social Presence (Figure 7.4). Thus, the next research step should focus on investigating this issue.

Moreover, Social Presence is typically studied along social context, online communication and interactivity (Tu, 2002). These dimensions can only be captured in the context of an interaction. The importance of context is further suggested by the results of the study reported in Chapter 3, which highlights that context improves the recognition of emotional body language. Consistent with De Gelder (2010), it was found that congruent context increases the recognition rate of emotional body language. It would therefore be beneficial to investigate if, within the context of an interaction, the loss in believability affects Social Presence. The context of an interaction may affect the perceived believability of a character, reinforcing its behaviours and Social Presence by presenting it within a situation. However, the effect of incongruence or imperfections in the context and in the interaction could have a negative effect on interpretation and could also result in an Uncanny Effect.

Moreover, to fully grasp the problem described by the Uncanny Valley, one important area to look at is emotional reactions (Figure 7.4). Indeed, in order for a virtual environment to be effective and provide accurate transfer of training, the simulated experience should seem real and engaging to participants (see Section 1.2). This should include emotional reactions. Moreover, a sign of presence within a Virtual Environment is when users tend to react as if the situation was real (Slater, 2003). Thus, investigation on Social Presence, within the context of an interaction, should also provide an avenue for studying users' emotional reactions and whether they are affected by the Uncanny Valley.

### 7.6.2. Other Aspects of Realism

The effects of Realism on perception of emotional body language were investigated in terms of anatomy and body movements (Figure 7.1). This has lead to valuable contributions. However, these are not the only elements that need to be considered in order to fully understand the relationships between realism and believability.

A different relevant area of investigation to consider is vocal expressions (Figure 7.4). This is especially relevant to investigate whether believability as an effect on Social Presence (see Section 7.1). The level of realism of vocal expression and how it is perceived by users should be carefully investigated as it may also contribute to the Uncanny Valley. Moreover, this will be especially important for a social interaction with an animated character to occur. The principles of animation (Appendix A) suggest that the voice of a character needs to be consistent with the way it looks. However, this adds an extra layer of complexity as it requires believable synthetic voices.

Another important element that needs to be considered in order to investigate the perception of animated characters within the context of an interaction is facial expression. The realism of facial expressions has been considered in previous research (see Section 2.3 for further discussions) and it has been argued that it may contribute to the Uncanny Valley. However, most of the studies only considered static expressions and it is not evident that their findings hold when introducing movements (Qiao, Eglin, & Beck, 2011). Dynamic elements have been found to increase the interpretation of facial expression (Kamachi et al., 2001) and hence it should be taken into account in order to investigate whether realistic animated characters can induce Social Presence.

To be truly realistic, the animated characters would have to express emotions as humans do, through voice, facial expressions and body language (Kalat & Shiota, 2007), without any exaggeration. Ultimately, research on the perception of animated emotional expressions need to encompass all the modalities used by humans.

## 7.7. Final Comment

Investigating the influence of realism on the perception of animated emotional body language contributed significant new knowledge. The thesis shows that the recognition of an emotion is not affected by the nature of the character displaying it.

Moreover, a number of directions for further research were identified. Taken together with the contribution to knowledge, the thesis can inform the design of better animated characters for social task training.

# Bibliography


Aharoni, E., & Fridlund, A. J. (2006). Social reactions toward people vs. computers: How mere lables shape interactions. *Computers in human behavior, 23*(5), 2175-2189.

AMS. (2007). http://www.actorsmovementstudio.com/summer/course.htm. Retrieved 22/05/2007

Anolli, L., Vescovo, A., Agliati, A., Mantovani, F., & Zurloni, V. (2006). Simulation-based training of communication and emotional competence for improvement of physician-patient relationship. *Annual review of Cybertherapy and telemedicine, 4*, 79-86.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717-746.

Aylett, R., Paiva, A., Woods, S., Hall, L., & Zoll, C. (2005). Expressive Characters in Anti-Bullying Education. *publication13280*.

Balda, K. (2008). Full of character. *3D world, February 2008*, 64-67.

Banos, R. M., Botella, C., Alcaniz, M., Liano, V., Guerrero, B., & Rey, B. (2004). Immersion and emotion: Their impact on the sense of presence. *Cyberpsychology & Behavior, 7*(6), 734-741.

Bartle, R. A. (2003). *Designing Virtual Worlds*. Berkeley, California: New Riders.

Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2007). *Is the Uncanny Valley an Uncanny Cliff?* Paper presented at the Proceedings of the 16 th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2007.

Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM, 37*(7), 122-125.

Beck, A., Cañamero, L., & Bard, K. (2010). *Towards an affect space for robots to display emotional body language.* Paper presented at the Ro-Man 2010, Viareggio, Italy.

Beck, A., Hiolle, A., Mazel, A., & Cañamero, L. (2010). *Interpretation of emotional body language displayed by robots*. Paper presented at the Proceedings of the 3rd international workshop on Affective interaction in natural environments.

Bickmore, T., Gruber, A., & Picard, R. (2005). Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Educ Couns, 59*(1), 21-30.

Brenton, H., Gillies, M., Ballin, D., & Chatting, D. (2005). The Uncanny Valley: does it exist and is it related to presence ? *Presence-connect*, *http://www.presence-research.org/*.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition, 83*(1), 1-29.

Cassata, D. M. (1983). Physician interviewing and counselling training. [Article]. *International Journal for the Advancement of Counselling, 6*(4), 297-305.

Cassell, J. (2000). Nudge nudge wink wink: elements of face-to-face conversation. In J. Cassell, T. W. Bickmore, L. Campbell, H. Vilhjalmsson & H. Yan (Eds.), *Embodied Conversational Agents* (pp. 1-27). Cambridge: MIT Press.

Cañamero, L. (2006). Did garbo care about the uncanny valley? commentary to k.f. macdorman and h. ishiguro, "the uncanny advantage of using androids in cognitive and social science research". *Interaction Studies, 7*(3), 355-359.

Chaminade, T., Hodgins, J., & Kawato, M. (2007). Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience, 2*(3), 206-216.

Chapman, N., & Chapman, J. (2004). *Digital multimedia*: John Wiley & Sons, Ltd.

Choi, K.-S., Soo, S., & Chung, F.-L. (2009). A virtual training simulator for learning cataract surgery with phacoemulsification. *Computers in Biology and Medicine, 39*(11), 1020-1031.

Coleridge, S. T. (1817). *Biographia literaria*.

Dantzer, R. (2005). *Les Emotions*. Paris: Presses universitaires de France.

Dargis, M. (2007). Beowulf (2007). *New York Times* Retrieved January 20, 2008

Darken, R. P., & Banker, W. P. (1998). *Navigating in natural environments: a virtual environment training tranfer study*. Paper presented at the Virtual Reality Annual International Symposium.

Dautenhahn, K., & Hurford, J. (2006). Interaction studies. *7*(3).

Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Mirza, N. A., et al.(2009). KASPAR – a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics, 6*(3), 369 - 397.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature reviews Neuro-Science, 7*(3), 242-249.

De Silva, P. R., & Bianchi-Berthouze, N. (2004). Modeling human affective postures: an information theoretic characterization of posture features. *Computer Animation and Virtual Worlds, 15*(3-4), 269-276.

den Stock, J. V., Grèzes, J., & de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research, 1242*, 185-190.

den Stock, J. V., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7*(3), 487-494.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3/4), 169-200.

Grantcharov, T. P., Kristiansen, V. B., Bendix, J., Bardram, L., & Funch-Jensen, P. (2003). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery, 91*(2), 146-150.

Gratch, J., Mao, W., & Marsella, S. (2006). Modeling social emotions and social attributions. In R.Sun (Ed.), *Cognition and Multi-Agent interaction: Extending cognitive modeling to social simulation*: Cambridge university press.

Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K., & Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *Human-Computer Studies, 67*(10), 842-849.

Hall, J. A., Harrigan, J. A., & Rosenthal, R. (1995). Nonverbal Behavior in Clinician Patient Interaction. *Applied & Preventive Psychology, 4*(1), 21-37.

Hamblin, C. J. (2005). *Transfer of training from virtual reality environments*. Wichita State University, Wichita.

Hanson, D. (2006). *Exploring the aesthetic range for humanoid robots*. Paper presented at the cognitive science (CogSci 2006) workshop on android science.

Hanson, D., Olney, A., Pereira, I. A., & Zielke, M. (2005). *Upending the uncanny valley*. Paper presented at the *A*merican association for artificial intelligence (AAII) conference.

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence: Teleoperators and virtual environments, 1*(2), 262 - 271.

Ho, C.-C., MacDorman, K. F., & Pramono, Z. A. D. (2008). *Human emotion and the uncanny valley: A GLM, MDS, and ISOMAP analysis of robot video ratings*. Paper presented at the International Conference on Human-Robot Interaction.

Huang, M. P., & Alessi, N. E. (1999, 1999). *Presence as an emotional experience*. Paper presented at the Medicine meets virtual reality, Amsterdam.

Jackson, P. (Writer). (2001). The lord of the rings.

Jarmon, L., Traphagan, T., Mayrath, M., & Trivedi, A. (2009). Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. *Computers & Education, 53*(1), 169-182.

Jerabek, I. (2001). Emotional Intelligence Test - 2nd Revision.

Kalat, J. W., & Shiota, M. N. (2007). *Emotion*. Belmont: Thomson Wadsworth.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception, 30*(7), 875-887.

Kleinsmith, A., Bianchi-Berthouze, N., & Steed, A. (2011). Automatic Recognition of Non-Acted Affective Postures. *IEEE Transactions on Systems, Man, and Cybernetics Part B*.

Kleinsmith, A., De Silva, P. R., & Bianchi-Berthouze, N. (2006). Cross-cultural differences in recognizing affect from body posture. *Interacting with Computers, 18*(6), 1371-1389.

Koda, T. (1996). *Agents with faces: A study on the effects of personification of software agents*. Massachusetts Institute of Technology, Massachusetts.

Kolb, D. A. (1984). *Experiential Learning: experience as the source of learning and development*. Englewood Cliffs, N.J: Prentice-Hall.

Konkle-Parker, D. J., Cramer, C. K., & Hamill, C. (2002). Standardized patient training: a modality for teaching interviewing skills. *J Contin Educ Nurs, 33*(5), 225-230.

Kramer, N. C. (2007). Nonverbal communication. In J. J. Blascovitch & C. R. Hartel (Eds.), *Human behavior in military contexts*.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills and London: Sage publications.

Lane, J. L., & Gottlieb, R. P. (2004). Improving the interviewing and self-assessment skills of medical students: Is it time to readopt videotaping as an educational tool? [Article]. *Ambulatory Pediatrics, 4*(3), 244-248.

Lee, K. M. (2004). *Presence, Explicated*. Paper presented at the Annual meeting of the International Communication Association.

MacDorman, K. F., Green, R. D., Ho, C. C., & Koch, C. (2009). Too real for comfort: Uncanny responses to computer generated faces. *Computers in human behavior, 25*(3), 695-710.

MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in social and cognitive science research. *Interaction studies, 7*(3), 297-337.

Mantovani, F., & Castelnuovo, G. (2003). Sense of presence in virtual training: Enhancing skills acquisition and transfer of knowledge through learning experience in virtual environments. In G. Riva, F. Davide & W. A. IJsselsteijn (Eds.), *Being there: Concepts, effects and measurement of user presence in synthetic environments*. Amsterdam: Ios Press.

Martin, J.-C., Niewiadomski, R., Devillers, L., Buisine, S., & Pelachaud, C. (2006). Multimodal complex emotions: Gesture expressivity and blended facial expressions. *International Journal of Humanoid Robotics, Special Edition "Achieving Human-Like Qualities in Interactive Virtual and Physical Humanoids", 3*(3), 269-292.

McBreen, H. M., & Mervyn, A. J. (2001). Evaluating Humanoid Synthetic Agents in E-Retail

Applications. *Transactions on systems, man, and cybernetics-Part A: Systems and Humans, 31*(5), 394-405.

McDonnell, R., Jorg, S., McHugh, J., Newell, F., & O'Sullivan, C. (2008). *Evaluating the emotional content of human motions on real and virtual characters*. Paper presented at the 5th symposium on Applied perception in graphic and visualization, Los Angeles, California.

Mori, M. (1970). Bukimi no tani [The un-canny valley] . *Energy, 7*, 33-35.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103.

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism of users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoper. Virtual Environ., 12*(5), 481-494.

Nunez, D. (2004). *How is presence in non-immersive, non-realistic virtual environments possible?* Paper presented at the Proceedings of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africa.

O'Sullivan, P., Chao, S., Russell, M., Levine, S., & Fabiny, A. (2008). Development and implementation of an objective structured clinical examination to provide formative feedback on communication and interpersonal skills in geriatric training. [; Research Support, U.S. Gov't, P.H.S.]. *J Am Geriatr Soc, 56*(9), 1730-1735.

Parsons, T. D., Kenny, P., Ntuen, C. A., Pataki, C. S., Pato, M. T., Rizzo, A. A., et al.(2008). *Objective structured clinical interview training using a virtual human patient*. Paper presented at the Medicine meets virtual reality (MMVR 16).

Pelachaud, C. (2009). Modelling multimodal expression of emotion in a virtual agent. *Philos Trans R Soc Lond B Biol Sci, 364*(1535), 3539-3548.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition, 82*(2), 51-61.

Pratt A, J., Hauser, K., Ugray, Z., & Patterson, O. (2007). Looking at human-computer interface design: Effects of ethnicity in computer agents. *Interacting with Computers, 19*(4), 512-523.

Prendinger, H., Ma, C., & Ishizuka, M. (2007). Eyes movements as indices for the utility of life-like agents: a pilot study. *Interacting with Computers, 19*(2), 281-292.

Qiao, S., Eglin, R., & Beck, A. (2011). *Audience perception of computer generated human behavior*. Paper presented at the CGAT 2011.

Raij, A. B., Johnsen, K., Dickerson, R. F., Lok, B. C., Cohen, M. S., Duerson, M., et al.(2007). Comparing interpersonal interactions with a virtual human to those with a real human. *Ieee Transactions on Visualization and Computer Graphics, 13*(3), 443-457.

Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television and new media like real people and places*: Cambridge University Press.

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., et al.(2007). Affective interactions using virtual reality: The link between presence and emotions. *Cyberpsychology & Behavior, 10*(1), 45-56.

Rizzolatti, G., & Fabbri-Destro, M. (2009). The Mirror neural system. In Wiley (Ed.), *Handbook of Neuroscience for the Behavioral Sciences*.

Robins, B., & Dautenhahn, K. (2007). Encouraging social interaction skills in children with autism playing with robots: A case study evaluation of triadic interactions involving children with autism, other people (peers and adults) and a robotic toy. *Enfance, 59*(1), 72-81.

Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., & Penn, P. R. (2000). Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics, 43*(4), 494 - 511.

Sakaguchi, H., & Sakakibara, M. (Writer). (2001). Final Fantasy: The Spirits Within. Japan: Sony Pictures.

Sanchez, M. V., & Slater, M. (2005). From Presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*(4), 332-339.

Scherer, K. R. (1987). Toward a dynamic theory of emotion: The component process model of affective states. *Geneva Studies in Emotion and Communication, 1*, 1-98.

Scherer, K. R. (2005). What are emotions? And how can they me measured? *Social Science Information, 44*(4), 693-727.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., et al.(1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*(2), 167-177.

Seyama, J. i., & Nagayama, R. S. (2007). The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence, 6*(4), 337-351.

Slater, M. (2003). A Note on Presence Terminology. *Presence-connect, 3*.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., et al.(2006). A virtual repirse of the Stanley Milgram obedience experiments. *PLoS ONE, 1*(1).

Slater, M., & Steed, A. (2000). A Virtual Presence Counter. *Presence: Teleoper. Virtual Environ., 9*(5), 413-434.

Swartout, W., Gratch, J., Hill, R. W., Hovy, E., Marsella, S., Rickel, J., et al.(2006). Toward virtual humans. *AI Magazine, 27*(2), 96-108.

Takeuchi, A., & Nagao, K. (1993). *Communicative facial displays as a new conversational modality*. Paper presented at the Proceedings of the INTERCHI '93 conference on Human factors in computing systems.

Thomas, F., & Johnston, O. (1995). *The illusion of life*. New-York: Abbeville Press.

Triola, M., Feldman, H., Kalet, A. L., Zabar, S., Kachur, E. K., Gillespie, C., et al.(2006). A randomized Trial of teaching clinical skills using virtual and live standardized patients. *Journal of General Internal Medicine, 21*(5), 424-429.

Tu, C.-H. (2002). The Measurement of Social Presence in an Online Learning Environment. *International Journal on E-Learning, 1*(2), 34-45.

Van de Riet, W. A. C., Grèzes , J., & de Gelder, B. (2009). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Social Neuroscience, 4*(2).

Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., et al.(2006). *Building Expression into Virtual Characters* Paper presented at the Eurographics 2006, Vienna.

Vora, J., Nair, S., Gramopadhye, A. K., Duchowski, A. T., Melloy, B. J., & Kanki, B. (2002). Using virtual reality technology for aircraft visual inspection training: presence and comparison studies. *Applied Ergonomics, 33*(6), 559-570.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28*(6), 879-896.

Wexelblat, A. (1998). Don't Make That Face: A Report on Anthropomorphizing an Interface. *Report SS-98-02, AAAI Press*.

Winters, A. (2005). Perception of body posture and emotion: A question of methodology. *The new school psychology bulletin, 3*(2), 35-45.

Zemeckis, R. (Writer). (2004). The Polar Express. USA: Warner Bross.

Zemeckis, R. (Writer). (2007). Beowulf. USA: Paramount Pictures, Warner Bros.

Zhang, L., Gillies, M., Dhaliwal, K., Gower, A., Robertson, D., & Crabtree, B. (2009). E-Drama: Facilitating Online Role-play using an AI Actor and Emotionally Expressive Characters. *International Journal of Artificial Intelligence in Education, 19*(1), 5-38.

# APPENDIX A

For social task training, it would be necessary to populate the Virtual Environment with believable and engaging characters behaving appropriately. In animation, a believable character is a character that seems alive to its audience. Traditional animation has a long history of creating believable and engaging characters by applying a set of design principles. This set of principles, if appropriately used could be used to animate the characters needed for social task training. This could create the illusion of life, which in turn should induce social presence. Thus, the principles used in traditional animation should be considered to achieve this illusion of life.

## A.1. Issues with the Principles of Animation

**Figure A.1:** Screen shot from *Snow White (Disney 1937)*

In traditional animation, believability is achieved by creating an "illusion of life" (Thomas & Johnston, 1995) and thereby avoiding the drop in believability described by the Uncanny Valley. For instance, the animated characters in Snow White (Figure A.1) can be placed at the peak above the Uncanny Valley (Section 2.2). This is possible only if the character displays emotion (Bates, 1994).

Traditional animation has been successful in creating an "illusion of life" by respecting some fundamental principles, which hold the keys to creating believable and emotionally engaging characters (Balda, 2008). They have been successfully used since the nineteen thirties and were introduced by Disney studios. However, their application to highly realistic characters, such as the ones needed for social task training, needs careful consideration.

### A.1.1. Traditional animation principles that can be used to increase realism

The first principle, is called "Squash and Stretch" (Thomas & Johnston, 1995) and is a way to emphasise movements by physically modifying an object. This was based on the observations that anything made of living flesh tends to show considerable movements within the shape when moved (Thomas & Johnston, 1995), for instance muscles tend to stretch when in tension. These movements are necessary to create believable animations as they emphasise the inner life of a character. Moreover, this principle could be used to achieve a greater level of physical realism in characters anatomy. For example, muscles tend to bulge when moved. However, in traditional animation it is usually stylised or overreacted in order to add 'character' to the animation. However, this could be problematic for realistic animation, such as social task training, where movement must be accurate. Although some Squash and Stretch should be present e.g. muscle deformation, it must be implemented at an appropriate level, and not exaggerated as the principle implies.

The second animation principle is known as "Anticipation". It is used to prepare the audience to see what is about to happen before it does (Thomas & Johnston, 1995). The assumption is that an audience watching a scene will not be able to understand the actions presented unless it is prepared to see them (Thomas & Johnston, 1995). In traditional animation, it is accomplished by preceding each important action with movements or sounds that help viewers anticipate the forth-coming action. For instance, in the case of medical interview training, a character about to faint could give realistic cues associated with its physical condition, such as sweating, loss of colour etc. Again though the exaggerated must be avoided.

The third and fourth principles refer to the speed of movements. The third principle is "timing" and emphasises the importance of the speed at which movements occur (Thomas & Johnston, 1995). The fourth principle is referred as "ease in and out". It emphasises the importance of graduating the acceleration and deceleration of objects or movements (Thomas

& Johnston, 1995). The speed of movement is used in traditional animation to improve emotional clarity, making it easier for the audience to understand the emotional state of a character. This effect is supported by research in the psychology of emotion, which shows that if a character performs the same movements at different speed, their emotional content is interpreted differently. Pollick et al (2001), when looking at the perception of emotions conveyed by an arm performing the same gesture at different speeds found that the speed of movements is one of the dimensions that viewers use to interpret an emotion (Pollick, et al., 2001) i.e. determine the inner emotional state not the actual action. Therefore timing and acceleration are relevant to realistic characters, as they would be key to displaying emotions in an understandable and realistic way, rather than simply moving.

The fifth animation principle is called “arcs” (Thomas & Johnston, 1995) and was defined following real world observations. It suggests that in real life, a movement is very seldom completed in a straight line, as it tends to follow an arc. Although, this lacks empirical support, it seems that realistic characters should follow this principle in order to create believable movements.

The sixth principle emphasises the importance of secondary action (Thomas & Johnston, 1995) i.e. muscles budge as an arm moves. This is relevant for realistic characters as it support the believability of the character by creating an holistic experience. In traditional animation, secondary actions are created in a way that can be seen but without distracting from the main action. Secondary actions are present in real life too and may be important for social skills training as they give clues regarding the emotional state of a character, for instance, a character walking angrily could close his fists while walking. Therefore, secondary actions could be used to increase behavioural realism and hence provide trainees with training situations similar to the ones that occur in real life.

The seventh principle is called “follow through and overlapping action” (Thomas & Johnston, 1995). This principle is very similar to anticipation but takes place at the end of an action. For example, during a walk cycle, the hip leads, followed by the leg and then the foot. As the lead part stops, the lagging parts continue in motion. As with anticipation, it is relevant in situations where it happens in real life. When it comes to animating characters for training, it cannot be implemented fully as in traditional animation. Overlapping actions are actions that occur because of others. This could be used to emphasize the emotional state of a character. Overlapping actions seems relevant as well for realistic characters. They are likely to be expected by the viewers. This should improve the behavioural realism and the believability of the characters.

The eighth principle is "Appeal" (Thomas & Johnston, 1995), where the audience should feel engaged with the characters. This principle is important for social skills training and realistic characters must be designed in a way that will emotionally engage the trainees. This will increase social presence (Section 1.2) but also allow trainees to learn to experience their own emotional state and respond to a training scenario.

The last principle that can be applied, while still preserving behavioural realism, is "Personality" (Thomas & Johnston, 1995). In traditional animation, different characters have different personalities and behave in a way that is consistent with their personalities. Thus, different characters behave and react differently if put in the same context. The characters that would be animated for social skills training would need to careful consider this area. However, the role of personality is outside the scope of this thesis as it is concerned with the perception of the resultant emotion, not its cause (Section 1.6.1).

Thus, these principles can be applied to training characters in order to reach a greater level of realism. However, in traditional animation they are often used an unsubtle that tend to be stylised or overreacted, which makes it obvious to the viewer but is different to what occurs in real life. For example, in traditional animation, a character about to faint could give several unrealistic but obvious cues for the audience to anticipate the fainting e.g. spinning on one leg. In the case of social task training, movements, actions and postures (Section 1.4) would have to be realistic or else trainees would ultimately look for the wrong cues and underestimate the complexity of real life situations, resulting in poor transfer of training. Hence, a patient about to faint in a medical interview scenario should use these principles to reflect the condition of the patient being portrayed not attempt to have the action obvious.

However, applying these principles in a realistic way may result in the animated character falling into the Uncanny Valley as a high level of behavioural realism may have a negative effect on believability. This is because consistency is important but there are principles of animation that cannot be used for characters required for social task training.

### A.1.2. Traditional animation principles that cannot be applied to social task training

The tenth animation principle is “Exaggeration” (Thomas & Johnston, 1995). In traditional animation, emotion is conveyed through exaggerated gestures. However, in a training simulation, the display needs to be realistic, so that students do not learn to rely on “incorrect” behavioural clues. Using overacted displays of emotions may result in training students to acquire skills that would not apply in real life situations where the emotions is often displayed in more subtle ways. Exaggeration must also be avoided when applying the other principles (Section 2.2.1).

The eleventh principle is called “Straight ahead action and pose-to-pose action” (Thomas & Johnston, 1995). These are two distinct techniques for animation. Straight ahead animation means that the animator draws each individual frame of a sequence starting with the first one. Pose-to-pose animation means that the animator sets up key poses and that the animation in between is automatically generated. While these two techniques are relevant to traditional animation, the complexity of producing realistic animations makes it very difficult to apply one of these. Moreover, the movie industry has started using motion capture technology in order to produce highly realistic animation. For example, movies such as Final Fantasy (Sakaguchi & Sakakibara, 2001), Polar Express (Zemeckis, 2004) or Beowulf (Zemeckis, 2007) relied on motion capture technology rather than hand animations. However, it may be this change that is causing a drop in believability.

The twelfth principle is “Staging” (Thomas & Johnston, 1995). In traditional animation, an action is presented (staged) so that the audience can easily understand it. However, rather than being easy to understand, animated characters for social task training should be staged in a realistic manner, simulating the complexity of real life situations. However, ignoring this principle may make the scene look “upstaged” (Thomas & Johnston, 1995), resulting in the trainee not knowing what to look at because too many things are occurring at the same time and therefore may not feel engaged with the animations.

Moreover, the idea that the more realistic a character is the more believable it will be has been challenged by realistic characters from the movie industry, which have failed to create empathy from their audience. For instance, even the most hyper-realistic human animated characters using high fidelity Motion Capture technology, such as the ones used in the $28.1

million production of Beowulf (Zemeckis, 2007), are confronted with the Uncanny Valley, which results in missing the “spark of true life” (Dargis, 2007). Thus, even though investment is being put into building simulations, that include realistic characters that will display emotions realistically, it is not known yet what effects such a display will have on a viewer.

To be truly realistic, the animated characters would have to express emotions as humans do, through voice, facial expressions and body language (Kalat & Shiota, 2007), without any exaggeration. Moreover, it has long been accepted in traditional animation, that expressions must be captured throughout the whole body as well as the face (Thomas & Johnston, 1995). Traditional animation emphasises the importance of body expressions and some of the principles, including “arcs”, squash and stretch, secondary action” focus mostly on body expression. This suggests that for social task training, it would be important to investigate the display of realistic body language, similar to that displayed in real life.