Emotional Body Language Displayed by Artificial Agents

ARYEL BECK, STRI & School of Computer Science, University of Hertfordshire, UK
BRETT STEVENS, School of Creative Technologies, University of Portsmouth, UK
KIM A. BARD, Centre for Comparative and Evolutionary Psychology, University Of Portsmouth, UK
LOLA CAÑAMERO, School of Computer Science, University of Hertfordshire, UK

Complex and natural social interaction between artificial agents (computer generated or robotic) and humans necessitates the display of rich emotions in order to be believable, socially relevant and accepted, and to generate the natural emotional responses that humans show in the context of social interaction, such as engagement or empathy. Whereas some robots use faces to display (simplified) emotional expressions, for other robots such as Nao, body language is the best medium available given their inability to convey facial expressions. Displaying emotional body language that can be interpreted whilst interacting with the robot should significantly improve naturalness. This research investigates the creation of an Affect Space for the generation of emotional body language to be displayed by humanoid robots. To do so, three experiments investigating how emotional body language displayed by agents is interpreted were conducted.

The first experiment compared the interpretation of emotional body language displayed by humans and agents. The results showed that emotional body language displayed by an agent or a human is interpreted in a similar way in terms of recognition. Following these results, emotional key poses were extracted from an actor's performances and implemented in a Nao robot. The interpretation of these key poses was validated in a second study where it was found that participants were better than chance at interpreting the key poses displayed. Finally, an Affect Space was generated by blending key poses and validated in a third study.

Overall, these experiments confirmed that body language is an appropriate medium for robots to display emotions and suggest that an Affect Space for body expressions can be used to improve the expressiveness of humanoid robots.

Categories and Subject Descriptors General Terms: Experimentation, Human Factors.

Additional Key Words and Phrases: Human Computer Interactions, Human Robot Interactions, Emotional Body Language

1. INTRODUCTION

Complex and natural social interaction between humans and artificial agents (computer generated or robotic) necessitates the display of rich emotions in order to be believable, socially relevant and accepted, and to generate the

Authors' addresses: Aryel Beck, Adaptive Systems Research Group, STRI & School of Computer Science, University of Hertfordshire, AL10 9AB, UK. E-mail: a.beck2@herts.ac.uk;
 DOI 10.1145/1290002.1290003
http://doi.acm.org/10.1145/ 1290002.1290003

natural emotional responses that humans show in the context of social interaction, such as engagement or empathy. Whereas some robots use faces to display (simplified) emotional expressions, for other robots such as Nao, body language is the best medium available given their inability to convey facial expressions. This feature also offers researchers the rare opportunity to investigate how much emotion can be conveyed without taking facial expression into account.

Existing artificial agents express emotions using facial expressions, vocal intonation [Schröder 2001], body movements and postures. For instance, Greta is a virtual human that can display emotions using its face in combination with gestures [Pelachaud 2009]. A very different example is “FearNot”, which is a Virtual Environment for educating children on the issues of bullying [Aylett et al. 2005]. “FearNot” uses cartoonish characters with which children bond and sympathise while watching bullying situations from different perspectives. “FearNot” portray animated characters that use their face and body to express emotions and have demonstrated good learning outcomes [Enz et al. 2008]. For computer agents, other modalities have also been proposed including the use of lighting and shadows as a medium for artificial agents to express emotions [de Melo and Paiva 2008].

Expressive robots have also been successfully created. For instance the face of Kismet [Breazeal 2002] conveys emotions based on nine prototypical facial expressions that ‘blend’ together along three axes: Arousal, Valence and Stance. Arousal defines the level of energy. Valence specifies how positive or negative the stimulus is. Stance defines how approachable the stimulus is. This approach, based on the Circumplex model of emotions [Russell 1980], defines an Affect Space in which expressive behaviours span continuously across these three dimensions, creating a rich variety of expressions. Another example is the Kobian humanoid robot that can express emotions using its face in combination with body movements [Zecca et al. 2009].

Research has also focused on achieving responsive behaviours, especially for Virtual Humans. For instance, Gillies et al. [2008] have created a method to create responsive virtual humans that can generate their own expressions based on motion capture data [Gillies et al. 2008]. However, it is not possible to transfer this method directly onto robots as they cannot reproduce the movements captured by motion capture as smoothly as virtual humans or without falling over. Moreover, it is not evident that the same body language displayed by a human and by an agent (computer generated or robotic) would be interpreted in a similar way. For instance, Beck et al. [2009] found a drop in believability and recognition rate when comparing the same emotional body language displayed by an actor and a computer agent (“animated character”) animated using motion capture. These results could be due to the fact that the motion capture technology used for the animation does not capture micro gestures and secondary animations such as breathing and hand movements. Breathing is known to participate to the expression of specific emotions [Dantzer 2005]. Hand and finger movements have been found to contribute to the expression of certain emotions as well [Wallbott 1998]. These would not necessarily be a problem for an agent as these cues could be procedurally recreated and hence would not be missing. However, these differences could be due to the nature of the agent itself (artificial agent vs. biological character) and to the difference in physical realism between the actor and the animated character [Beck et al. 2008].

Indeed, it is not evident that expressions displayed by a human and by an agent are interpreted in a similar way. Encouraging results have however been found from a much weaker stimulus. Using restricted technology, it was found that humans tend to interact with computers as they do with real people [Nass and Moon 2000].

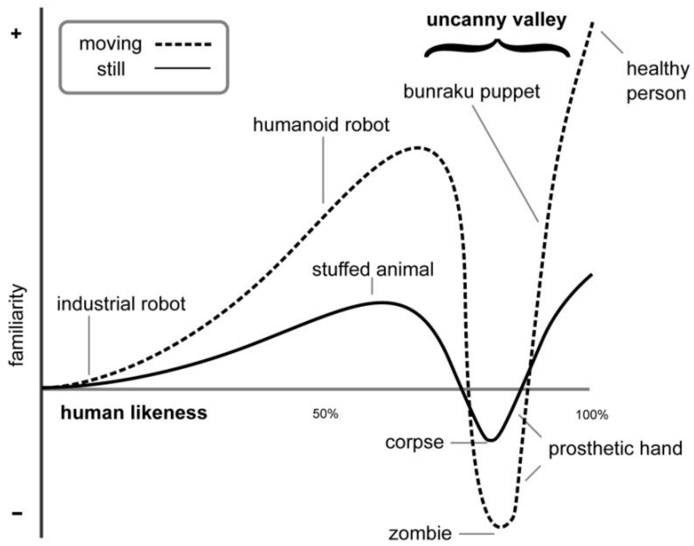

Fig. 1. The Uncanny Valley (adapted from Mori 1970)

However, it cannot be simply assumed that greater fidelity would improve the interaction. As agents become more visually realistic, they are confronted with the well-known Uncanny Valley problem [Mori 1970]. The Uncanny Valley (Figure 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]. In recent years, improvement in the field of animation technology has increased the level of visual realism that can be achieved. In the context of animation, visual realism is defined in term of physical realism (i.e. how similar to a human the character looks) and behavioural realism (i.e. how similar to a human the movements are). 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 as described by the Uncanny Valley. For example, the characters from "The Polar Express" [Zemeckis 2004] or "Final Fantasy" [Sakaguchi and Sakakibara 2001] have failed to be convincing [MacDorman et al. 2009]. Unrealistic characters, such as the ones used in traditional cel animation, were not confronted with this issue as they generated empathy. Further discussions on the concept of the Uncanny Valley can be found elsewhere [Dautenhahn and Hurford 2006].

However, the Uncanny Valley (Figure 1) is not grounded on systematic studies and its very existence is still subject to debate. Moreover, existing studies seem far too simple to fully explore the full complexity of the problem. The causes of it might not be straightforward, and might involve a complex combination of all sorts of contextual, cultural, social, and other factors [Cañamero 2006]. Moreover, part of the uncanny effect could be due to poor design or disappointing storylines. The Uncanny Valley may also result from issues with the display of body language or, as suggested by Brenton et al. [2005], from poor facial animation. Traditional animation which created a wide range of credible characters, avoiding the Uncanny Valley, highlights the importance of displaying appropriate emotional behaviours [Bates 1994; Thomas and 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; Kleinsmith et al. 2011; Kleinsmith, De Silva and Bianchi-Berthouze 2006]. This suggests that the Uncanny Valley could be due to poorly animated body language as well. This is supported by 'classical' animation which asserts that the expression must be captured throughout the whole body as well as the face [Thomas and Johnston 1995]. Even theatre upholds this principle, 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 2011] for example). Hence, an animated character or a robot displaying emotion realistically through the face and not through the rest of the body may look unnatural to a viewer as the Uncanny Valley may result from the body language being displayed as well. This possibility remains unexplored as most of the studies on the Uncanny Valley focused on facial expressions (see for instance [Hanson et al. 2005; MacDorman, Green, Ho and Koch 2009]). The Uncanny Valley threatens the implicit assumption in developing hyper-real characters; that viewers perceive and interpret a human's or agent's expressions in a similar way. 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.

Emotional cues that can be accurately interpreted are essentials in order to create believable characters [Bates 1994]. Although, it has been shown that emotional body language is an effective medium used by humans to convey emotions and that it can be used to automatically recognize emotions, it is not evident that the same gestures would be interpreted similarly by humans when displayed by an artificial agent. This is an important issue which has not been previously investigated. This is why the first experiment compared the interpretation of emotional body language displayed by humans and agents. The results showed that emotional body language displayed is interpreted in a similar way when displayed by agents or by humans. Following these results, emotional key poses were extracted from actor performances and implemented on a Nao robot. The interpretation of these key poses was validated in a second study were it was found that participants were better than chance at interpreting the key poses displayed. Finally, an affect space was generated by blending key poses and was validated in a third study.

2. COMPARING EMOTIONAL BODY LANGUAGE DISPLAYED BY HUMANS AND AGENTS

2.1. Design

The aim of this study was to compare the interpretation of body language displayed by humans and artificial agents. Existing results on the perception of emotions have shown that it is possible to identify emotions looking at body language only [Coulson 2004; de Gelder 2009; Kleinsmith, Bianchi-Berthouze and Steed 2011; Kleinsmith, De Silva and Bianchi-Berthouze 2006; Wallbott 1998]. However, the effect of changing the embodiment (i.e. human vs. artificial agent) has not been empirically tested. In order to ensure that the emotional body language displayed is similar across conditions, motion capture technology was used to record movements. Since it would be impossible to display the recorded movements on a robotic platform without unwanted modifications (body dimension and weight repartition is different), animation was used to avoid confounding the results.

Table I. Experiment conditions. The effects of *Character Type* and *Action Style* were tested within subjects and *Frame Rate* between subjects.

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

The Independent Variables were defined to test predictions based on the Uncanny Valley model. Realism was divided in two distinct categories, the level of physical realism (i.e. appearance) 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 with an anatomically accurate form. The most realistic character used was an actor video recorded, followed by a realistic computer generated character and a simplified one (Actor Vs. Realistic Character Vs. Simplified Character). The fact that traditional animation, such as Disney [Thomas and Johnston 1995], usually uses stylised and unrealistic displays of emotions, which are adapted to the physical appearance of the characters, may have an effect at different levels of physical realism. This could affect the perception of the animated display (participants on seeing a simplified agent 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 from the ones that could be displayed in the real world through to the ones that can be found 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 I).

It is unclear whether the movements that motion capture failed to record have an effect on how 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 was done by reducing the frame rate from twenty-five frames per second to twelve frames per second (12FPS Vs 25FPS). Twelve frames per second was chosen as it is usually considered just enough to maintain a satisfactory illusion of continuous motion [Chapman and Chapman 2004], therefore creating a condition in which body movements are slightly jerky but the illusion of motion is maintained (*Frame Rate*).

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 (Section 1). 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 (Section 1). Believability and naturalness are very important in animation as they contribute to the 'illusion of life' [Thomas and 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 (Figure1). 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.

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.

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*?

2.2. Participants

40 Participants were recruited, mostly members of staff and 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 two groups (12Frames/Second vs. 25 Frames/Second) (Table 1). The twenty participants (12 Females and 8 males) in the 25 Frames/Second 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 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.

2.3. Apparatus

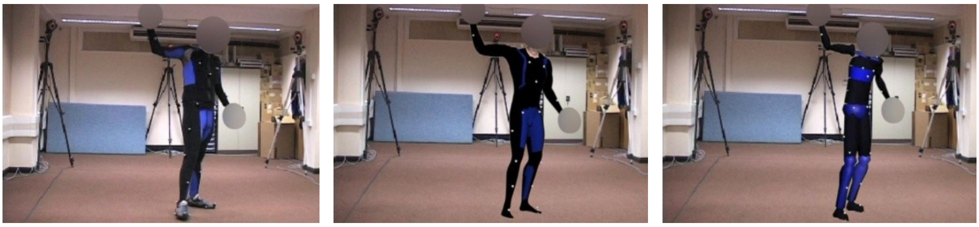

Fig. 2. Screen shots of the actor, of the realistic character and of the simplified character.

A professional actor and a professional director were hired to generate the performance of body movements in different emotional states. The actor performed ten emotions: Anger, disgust, shame, fear, sadness, surprise, relief, happiness, pride and excitement. Each emotion was performed in two different ways, an ordinary version for which the actor tried to act naturally and an exaggerated or stylised one closer to the style of traditional animation. He also performed a 'neutral' state. To ensure equivalency across conditions, the actor was video recorded (Figure2) and motion captured simultaneously. The videos were recorded using a Sony PD170P. Motion capture data was recorded using an eight-camera VICON motion capture system. The motion capture data was then used to animate two characters (Figure2) so that they displayed the same body language. The faces and hands of the actor and of the animated characters were pixelated, so that this source of information was removed along with the possible uncanny effect that may come from poor facial or poor hand animations. 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 (Figure2), were sized to look similar (faces were not visible) and were put in the same context (Figure2). Thus, sixty-three videos were created: (10 emotions x 2 Style (Natural and Stylised) + 1 neutral state) x 3 Character Types (Actor, Realistic Character and Simplified Character).

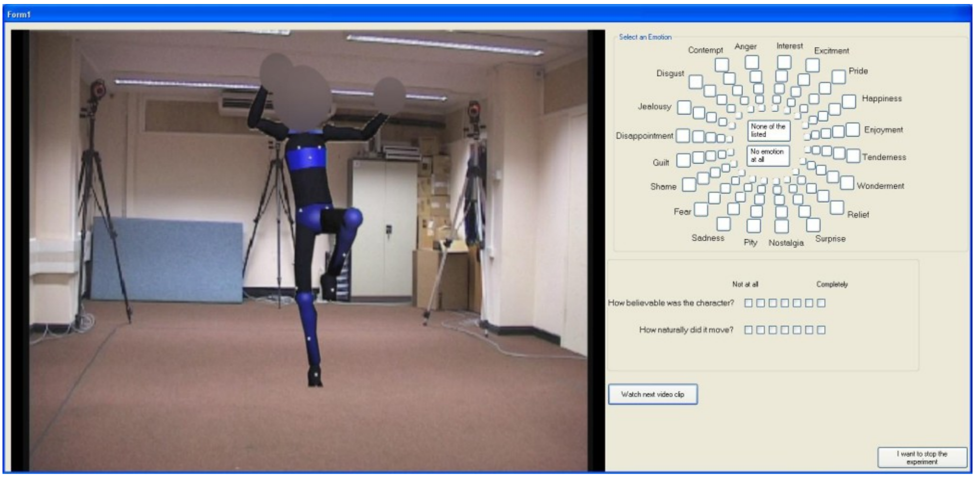

Fig. 3. Screen shot of the application

The videos, the animations and the questionnaire were displayed on a 5m x 2.5m rear projection screen. To record participants' answers, the material was embedded into a custom-made interface, which was used for displaying the video clip at life size as well as recording the participants' answers (Figure3).

2.4. Materials

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

For each emotional 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 in Video Games* 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.

Emotional Intelligence was measured using the 33-item emotional intelligence scale [Schutte et al. 1998].

2.5. 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 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 GEW. 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.

3. RESULTS

3.1. Preliminary data validation

To test 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 data. There was no difference between groups for *Emotional Intelligence* ($F(1,38)=0.01$, $p=0.92$), *Experience in Video Games* ($F(1,38)=0.00$, $p=1.00$), *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.

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

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

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

3.2. Overview of the Experimental Effects

The 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*.

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*. *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 $\eta^2=0.01$). 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 were combined for the rest of the analysis, effectively eliminating *Frame Rate* as a condition.

The effects of Character Type were assessed with Repeated Measures ANOVAs on each dependent variable; 3 (*Character Type*)) x 2 (*Action Style*). Assumptions of normality, homogeneity of variance and sphericity were met.

Character Type had no effect on *Correct Identification* ($F(1,38)=0.84$, $p=0.44$, Partial $\eta^2=0.02$).

Character Type had 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 4).

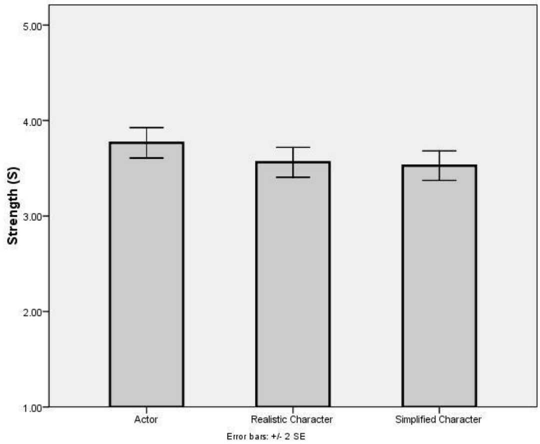

Fig. 4. The overall *Strength* by *Character Type* (Means + 2*Standard Error)

Character Type 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 ($p<0.01$) (Figure5).

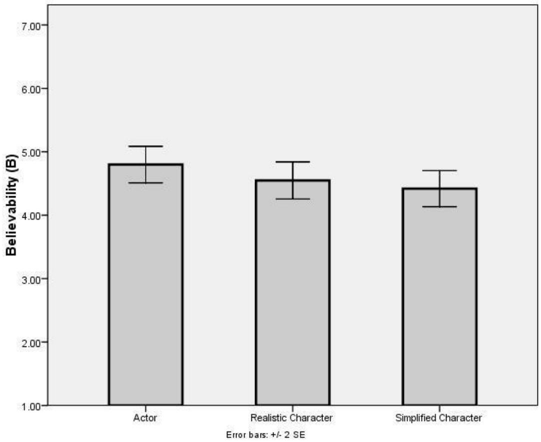

Fig. 5. The overall *Believability* by *Character Type* (Means + 2*Standard Error)

Character Type 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 ($p<0.05$) (Figure6).

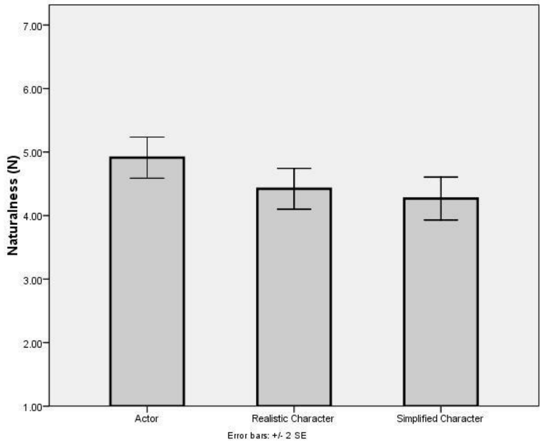

Fig. 6. The overall *Naturalness* by *Character Type* (Means + 2*Standard Error)

The general patterns from the investigation 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* (Figure4), on *Believability* (Figure5) and on *Naturalness* (Figure6).

Moreover, the recognition rate varied greatly depending on the emotions displayed (Table 2). This might have resulted in the lack of effect of *Character Type* on *Correct Identification*. Therefore, this was further investigated.

3.3. Effect of Character Type on Correct Identification

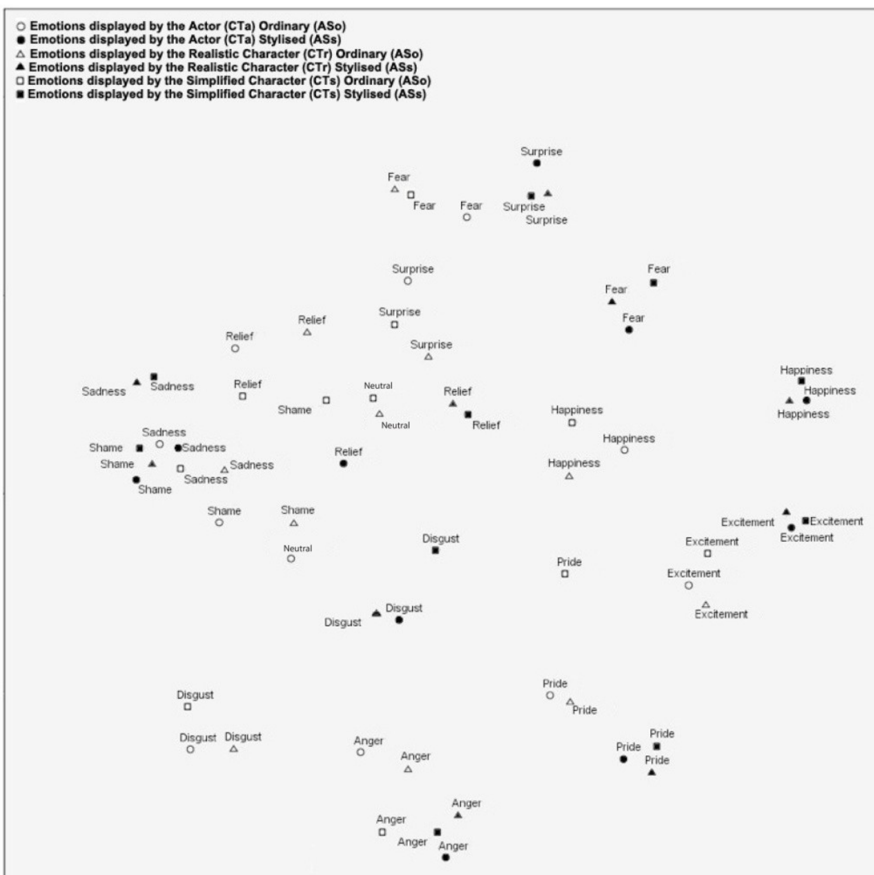

Fig. 7. The psychological space of interpretation of emotions display obtained by Multidimensional Scaling.

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).

Multidimensional scaling is a systematic procedure for obtaining a spatial representation consisting of a geometric configuration of points [Kruskal and Wish 1978]. Each point in the configuration (Figure7) corresponds to one of the 63 videos (10 emotions x 2 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 videos clips were generated by counting the percentage of interpretations each video clip had in common with the others. Video clips that were interpreted similarly had smaller distances from one another than those that were not. Videos clips were labelled by intended emotion (rather than by interpretation). The resulting solution is a unique two-dimensional psychological space (Figure7).

The clustering of videos clips in the psychological space (Figure7) show that the intended emotions cluster together independently from the character presented. This suggests that the lack of general significant differences 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.

3.4. Effects of Action Style

The effects of *Action Style* were assessed with 3 (Character Type) x 2 (Action Style) Repeated Measures ANOVAs on each dependent variable to determine the general patterns after eliminating the between subjects Frame Rate condition. Assumptions of normality, homogeneity of variance and sphericity were met.

Action Style had a significant effect on *Correct Identification* ($F(1,38)=78.73$, $p<0.01$, Partial $\eta^2=0.67$). This effect is also reflected in the psychological space (Fig 7) and confirms the accuracy of the figure. Moreover, *Action Style* had a significant effect on *Strength* ($F(1,38)=134,54$, $p<0.01$, Partial $\eta^2=0.77$) as well as on *Believability* ($F(1,38)=7.03$, $p<0.05$, Partial $\eta^2=0.15$). *Action Style* had no effect on *Naturalness* ($F(1,38)=0.42$, $p=0.52$, Partial $\eta^2=0.01$).

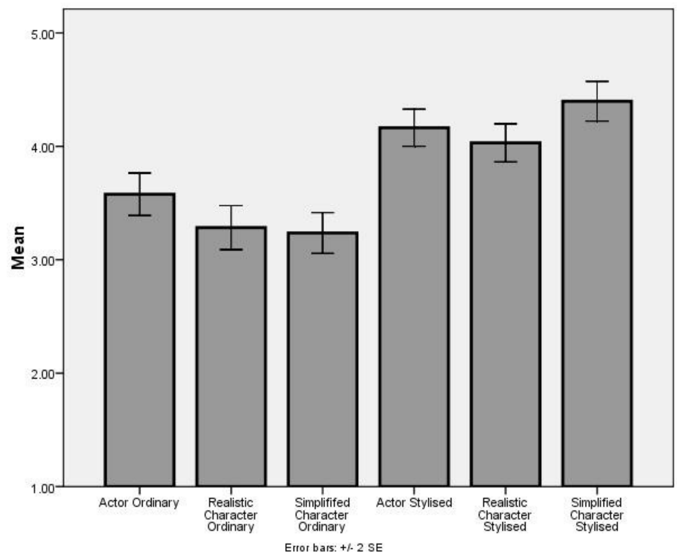

Fig. 8. The overall *Strength* for the two styles for each character (Means + 2*Standard Error)

There was no interaction between *Action Sty*le and *Character Type* for *Correct Identification* (F(1,38)=0.20, p=0.82, Partial η^2=0.01). However, as suggested by the hypotheses (Section 2.1), the results showed significant interactions between *Character Type* and *Action Style* for Strength (F(1,38)= 27.10, p<0.01, Partial η^2=0.41). In other words, the effect of *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 η^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). *Strength* was similar between the realistic character and the simplified character (p<0.45) (Figure8). 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 η^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) (Figure8).

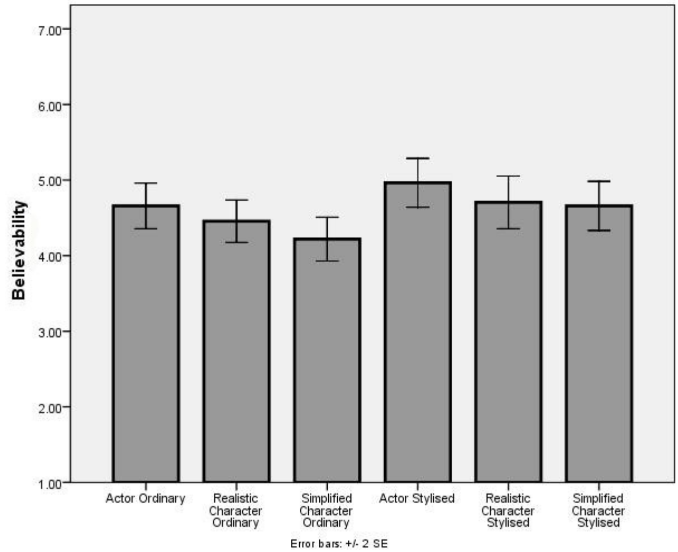

Fig. 9. The overall *Believability* for the two styles for each character (Means + 2*Standard Error)

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 ($p<0.01$) (Figure9). For the stylised displays, there was a significant difference in *Believability* 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* between the realistic character and the simplified character ($p=0.60$) (Figure9).

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

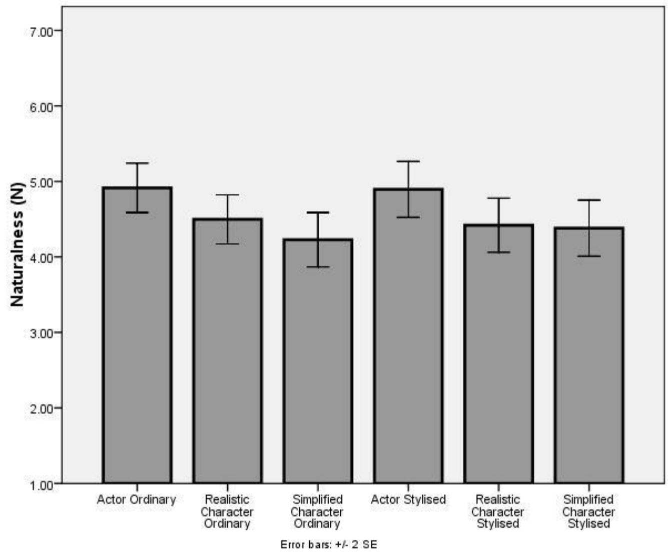

Fig. 10. The overall *Naturalness* for the two styles for each character (Means + 2*Standard Error)

There was a significant 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 ($p<0.01$) (Figure10). 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 ($p=0.69$) (Figure10).

3.5. 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 ($r(38)=0.15$, $p=0.55$).

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).

Interestingly, in the post hoc questionnaire, 9 out of 40 participants reported that during the study both the simplified character and the realistic character were actors whilst 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, with 24 participants thinking that the realistic character was real.

4. DISCUSSION

4.1. *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 II) even though faces and hands were hidden (Figure2). It can therefore be concluded that it is possible to correctly interpret the emotions that were displayed through watching body language alone.

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. The overall similarity of identification for the different characters was confirmed by the clustering of the emotions displayed in the psychological space (Figure7) resulting from the multi dimensional scaling. In Figure7, 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. Moreover, Figure7 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 existing results [Beck et al. 2009; McDonnell et al. 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 Character Types and the simplification of the level of physical realism has not degraded the emotional 'channel' of information for most of the displays.

4.2. *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. Moreover, there was no difference in emotional *Strength* between the realistic character and the simplified character. 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 could 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, it could explain why these emotions were still correctly identified but were perceived as less emotional when displayed by the animated characters. Therefore, the missing micro gestures and secondary animations did have an effect on the perceived emotional *Strength*. 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.

4.3. *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. 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*.

4.4. *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. This suggests that the body of the realistic character was 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. This is consistent with the animation principle which states that animated characters should move accordingly to the way they look [Thomas and Johnston 1995]. The lack of realism of the simplified character makes it look more 'cartoony' (Figure 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 artificial agents look and the way they move.

There was no effect of *Action Style* on *Naturalness*. This result and the fact that *Character Type* had an effect suggest that perceived *Naturalness* is independent from the body language displayed. It suggests that the important factor is the physical appearance of the character displaying the body language and not the actual body language itself.

4.5. 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 and test the potential effect of the loss due to the motion capture technology on the perception of emotional body language. The results are 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.

4.6. 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. 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 Beck et al. [2009], 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 in 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 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 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 (Figure1) 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.

4.7. Application to robotics

Overall the identification of emotional body language was found to be unaffected either by degrading the videos or by changing the biological character to an artificial agent. If this can be extended to robotic agents, it would present a strong advantage for many humanoid robots that do not have the ability to display facial expressions, such as Nao [Aldebaran 2010], and for which the body is the main medium available to express emotions visually. The results of this study suggest that users would interpret emotional body language displayed by a robotic agent accurately and that the expression would not be affected by an eventual lack of smoothness of its movements. Moreover, the study suggests that stylized displays of emotions would be more appropriate as it should be interpreted as more emotional and more believable on 'low realism' robots. Emotional expression is often implemented on robotic platforms and used to facilitate human-robot interactions. The emotional feedback provided by a robot can for example be intuitively used by humans to establish whether or not an interaction was successful [Hiolle et al. 2010].

5. CREATING AN AFFECT SPACE FOR ROBOTIC BODY LANGUAGE

The results reported in the previous section show that physical realism is not an issue, and contribute to a corpus of research that suggests that body language would be an appropriate modality for robots to display emotions. More precisely, various researchers have shown that emotions can be identified without speech or facial expressions [Atkinson et al. 2004; Beck, Stevens and Bard 2009; Grammer et al. 2004] and that distinctive patterns of movements and postural behaviours are associated with specific emotions [De Silva and Bianchi-Berthouze 2004; Kleinsmith, Bianchi-Berthouze and Steed 2011; Kleinsmith, De Silva and Bianchi-Berthouze 2006; Wallbott 1998]. However, the contribution of each body part to the emotional expressions has not been clearly established. This lack of knowledge regarding emotional body language makes it difficult to create systems capable of generating their own emotional expressions. For example, it is not currently possible to predict the effect that moving specific joints will have on the expressiveness of the body language being displayed. In other words, there is no equivalent to the Facial Action Coding system [Ekman et al. 2002] for the body, with very few attempts to study bodily expression in a systematic way [Coulson 2004; Kleinsmith, Bianchi-Berthouze and Steed 2011; Kleinsmith, De Silva and Bianchi-Berthouze 2006]. This is an important issue in the psychology of emotion that also has consequences for the creation of expressive artificial systems. Humanoid robots such as Nao [Aldebaran 2010] present an ideal platform for investigating emotional body language as they can display rich body language without varying facial expressions. Using such platform, it is possible to study body language independently from other modalities.

However, using an agent instead of a human may have an effect on the interpretation of emotional body language.

In animation, a well-established method for creating convincing and believable displays consists in starting from the creation of expressive key poses rather than body language in motion [Thomas and Johnston 1995]. In the context of emotional body language, a key pose is a static posture modelled so that it clearly describes the emotion displayed. Once the key poses are realized in robotic platforms, they can be used to drive the expressive animated behaviours. This method of creation was selected for the robot because it is possible to independently manipulate the position of joints and test the effects on the expressiveness of the key poses. If expressive key poses can be automatically generated, they can then be used to drive the expressive behaviours of the robot.

It may be possible to create a full range of expressions by blending a minimal set of body postures along a continuous model of emotions. Such an approach has been successfully implemented for facial expression. For example, the robot Feelix conveys emotions based on six basic prototypical facial expressions that can be merged to show more subtle emotions by showing a basic emotion in the upper part of the face and a different one in the lower part [Canamero and Fredslund 2001]. The robot Kismet uses nine prototypical expressions that 'blend' together along a continuous model of emotions [Breazeal 2002].

The main objective of this research is to create an Affect Space for body language. As a starting point, the contribution of the position of important joints to the expressiveness of displayed body language was tested along the dimensions of continuous model of emotions. Arousal and Valence are considered as important components of emotions [Russell 1980]. Stance was selected to reflect the importance of proxemics (i.e. interpersonal and environmental space) during social interactions [Argyle 1988]. Once the contribution of the important joints to the emotional expression is established, it should be possible to create expressive behaviours that blend continuously across dimensions. Such knowledge could also be used as guidelines for creating emotional body movements.

Head movements and position has been reported to be an important component of emotional expressions [Castellano 2008; Davidson 1994; Roether et al. 2009; Schouwstra and Hoogstraten 1995]. Moreover, animation emphasizes the importance of creating strong silhouette [Thomas and Johnston 1995] and it is expected that manipulating the head position could considerably change a robot's silhouette. Since this could greatly enrich the expressions generated by an affect space, it was decided to focus on the effect of altering the head position on a robotic platform before developing a 'complete' affect space.

6. EFFECT OF HEAD POSITION ON THE INTERPRETATION OF EMOTIONAL BODY LANGUAGE

Another experiment using a robotic platform was designed to extend the results of the first study and to test the effect of moving the head up or down in a range of different key poses. The position of the head was chosen as a starting point due to the broad range of emotions that it can convey and therefore its high potential to increase the effectiveness of the emotional expressions generated by an affect space. The platform chosen for this experiment was the robot Nao [Aldebaran 2010].

The experiment used a within-subjects design with two independent variables: *Emotion Displayed* and *Head Position*. The effect of changing the *Head Position* may vary depending on the position of other joints. In other words, the effect of moving the head up or down may differ depending on the emotion being displayed. Therefore, it was tested across six emotions (*Emotion Displayed*): Anger, Sadness, Fear, Pride, Happiness and Excitement. These emotions were chosen because their recognition rate was high in the first experiment and they ought to be distant from each other along the three axes of the Affect Space (they vary in Valence, Arousal and Stance), widening the area explored. For instance, it was expected that Happiness and Excitement would differ along the Arousal dimension.

Head position had three levels (Up, Down, and Straight), defined as the head position relative to the chest. Four dependent variables were defined to explore the Affect Space: *Correct Identification*, *Arousal*, *Valence* and *Stance* (see Section 1 for definitions). *Correct Identification* was used to test whether or not it was possible for participants to interpret the emotion of the key poses.

The three main questions tested were:
(Q1) Is it possible for participants to correctly identify the key poses displayed by Nao?
(Q2) Is the effect of moving the head similar across all the key poses? In other words, is the contribution of head position independent from the rest of the expression?
(Q3) What is the effect of changing the head position on the interpretation and perceived place of a key pose in the Affect Space?

6.1. Participants

26 Participants were recruited, mostly students of the University of Portsmouth (9 females and 17 males) ranging in age from 18 to 55 (M=29.31, SD=11.93). Participants were entered in a raffle to win an iPhone in exchange for participation.

6.2. Apparatus

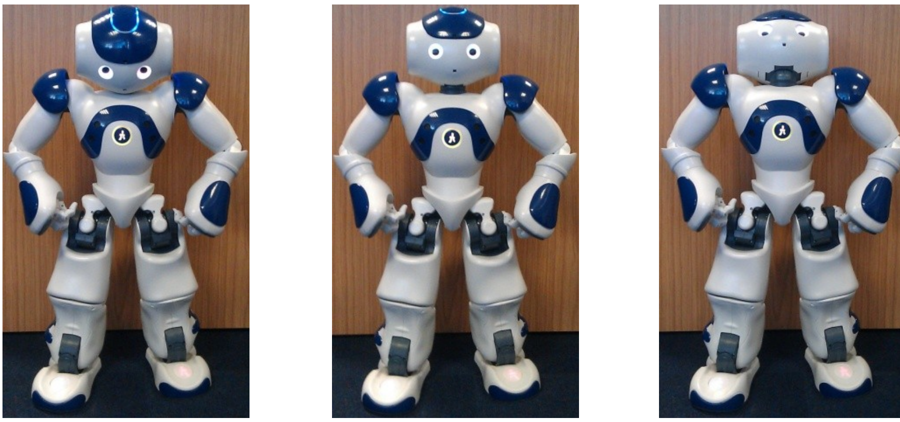

Fig. 11. The pride key pose in the three conditions

The six key poses were constructed by using the motion captured performances from the previous study. For each emotion, an expressive key pose was selected from the performance, based on its expressivity and on the likelihood of displaying it in the robot. Each joint of the robot was carefully positioned to match the original pose (Figure 11).

The experimental poses were then generated by systematically altering the head positions for each of the 6 key poses (Figure 11). For Head Position-Down, the head was rotated vertically all the way down. For Head Position - Up, the head was moved vertically completely up. For Head Position-straight, the head was aligned with the chest. This resulted in 18 poses (6 Emotion Displayed by 3 Head Position).

6.3. Preliminary Data Validation

A small-scale pilot study (with 7 participants) was conducted to test the quality of the six key poses and verify that the emotional labels are associated with the key poses. They were all recognized above chance levels when displayed by Nao (Anger 2/7, Sadness 4/7, Fear 6/7, Pride 7/7, Happiness 2/7, Excitement 2/7. Chance level would be 1/6).

6.4. Procedure

The same experimenter tested all participants individually. Once each participant had given consent at the beginning of their session, they were given some explanation regarding the questionnaire that they were expected to answer and were instructed to "imagine that the robot is reacting to something". In this context, Valence was defined as the extent to which this 'something' was positive or negative. Stance was defined as the extent to which the robot would tend to move toward this 'something' or to avoid it. Finally, Arousal was defined as the level of energy (low to high energy).

After confirming that they understood all the questions, participants watched and assessed the 18 poses. Each pose was displayed only once in a randomized order different for each participant. For each pose, participants were asked to assign an emotion label chosen from a list of six emotions. The list was comprised of Anger, Sadness, Fear, Pride, Happiness and Excitement. Ratings of Valence, Arousal, and Stance were completed by participants on a 10-point Likert scale. When all the poses were assessed, participants were fully debriefed. The sessions lasted approximately 30 minutes.

7. RESULTS AND DISCUSSION

In this section, only the relevant results are presented. A more complete analysis can be found in [Beck et al. 2010].

7.1. Is it possible for participants to correctly identify the key poses displayed in a robotic platform?

Table III. Percentage of participants that identified correctly the emotion at least once (chance level at 42%).

Anger	Sadness	Fear	Pride	Happiness	Excitement
88%	85%	92%	88%	73%	73%

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 (Table III). These recognition rates were obtained using static key poses only. Moreover, the relatively low recognition rates for Happiness and Excitement were mainly due to these two emotions being mistaken for one another (Table III).

The recognition rates clearly confirm the results from the first study as it is possible to interpret these emotions when displayed by a humanoid robot and that neither the lack of facial expression or the low level of realism is a barrier to expressing emotions. Although, the original postures used in this study were acted, the fact that it is possible to accurately identify them when displayed by a robot suggests that they could be used to improve robots social skills. This is important as social robots need to be able to express their internal states in order to interact with humans in a natural and intuitive way [Jamy et al. 2009].

7.2. What is the effect of moving the head on the interpretation and perceived place of a key pose in the Affect Space?

Repeated Measures ANOVAs (6 *Emotion Displayed* x 3 *Head Position*) were conducted on each dependent variable. Assumptions of normality, homogeneity of variance and sphericity were met.

Head Position had no significant main effect on *Correct Identification* ($F(2,48)=2.04$, $p=0.14$, Partial $\eta^2=0.08$). However, there was a significant interaction between *Emotion Displayed* and *Head Position* ($F(10,240)=8.68$, $p<0.01$, Partial $\eta^2=0.26$). These results indicate that the effect of *Head*

Position on *Correct Identification* depended on the individual emotion being displayed. A detailed analysis can be found in [Beck, Cañamero and Bard 2010].

Head Position had a strong effect on the interpretation of the key poses being displayed. For instance, the anger display was interpreted as happiness/excitement by a majority of participants when the head was up. However, with the head in the straight position, a majority of participants interpreted it correctly. Similarly, the pride display with the head up was interpreted as pride whereas moving the head down changed the interpretation to anger for the majority of participants (Figure 11).

Fear was not affected by the change in *Head Position* and was correctly interpreted in all conditions. This could be due to participants interpreting the change in *Head Position* as an indicator to where the stimulus that causes the emotion would be located.

7.3. What is the effect of moving the head on the interpretation and perceived place of a key pose in the Affect Space?

Overall, moving the robot's head upward increased *Arousal*, *Valence* and *Stance* whereas moving it down decreased it. In other words, the key poses with the head up were perceived as more energetic, more positive and more approaching than the key poses with the head down. The straight position was in between these two extremes.

Maximum arousal, positive valence and approaching stance are expressed with the head completely up. At the other extreme, the lowest arousal, negative valence and avoiding stance are expressed with the head completely down. This was further suggested by the shift in interpretations. Moreover, the results of the straight position which was perceived in between the two others suggest that it will be possible to enrich the range of expressions by moving the head between the two extreme positions, expressing less extreme emotions.

These results can already be integrated in an automated expressive system. The robot could automatically change his head position to express changes in his internal Arousal, Valence and Stance.

7.4. Is the effect of moving the head similar across all the postures?

Repeated Measure ANOVAs showed no significant interactions between *Emotion Displayed* and *Head Position* for *Arousal* ($F(10,240)= 1.26$, $p=0.28$, Partial $\eta^2=0.05$). However, there was a significant interaction for *Valence* ($F(10,240)= 4.84$, $p<0.01$, Partial $\eta^2=0.17$) as well as for *Stance* ($F(10,240)=3.70$, $p<0.01$, Partial $\eta^2=0.13$).

These results indicate that the effects of changing the *Head Position* on *Valence* and *Stance* depend on the rest of the body position (Emotion

Displayed). Therefore, to investigate the interaction, repeated measure ANOVAs were carried out for each emotion.

There was no interaction between *Head Position* and *Emotion Displayed* for *Arousal*. This indicates that the effect of moving the *Head Position* on *Arousal* was consistent across the different conditions.

However, significant interactions were found between *Head Position* and *Emotion Displayed* for *Valence* and *Stance*. The detailed analysis showed that there was no change of direction [Beck, Cañamero and Bard 2010]. In other words, changing the *Head Position* had always the same effect on *Valence* and *Stance* or did not have an effect at all.

The lack of effect of changing the *Head Position* for some of the key poses suggests that a rich Affect Space could not be generated by changing only the head position. It will be necessary to consider the effect of moving the rest of the body to increase the range of expression. For instance, the interpretation of fear was not affected by the changes in *Head Position*. The results suggest that it would not be possible to increase or decrease the expression of fear using this key pose and moving only the head. However, it is possible to change the expression of the other five emotions tested by just changing the head position. For instance, while displaying sadness, the robot could move the head up in order to express a less negative feeling that is interpreted as anger or sadness (i.e. effectively increasing the valence, arousal and stance of the expression).

8. CONCLUSION OF STUDY 2

This study confirmed that body language can be successfully used by humanoid robots such as Nao to express the six emotions tested. The results also show that changing the head position affects the expressiveness of the key poses. It was found that moving the head down leads to decreased arousal (defines the level of energy), valence (defines whether a stimulus is positive or negative) and stance (defines whether a stimulus is approachable) whereas moving the head up increases these three dimensions. This suggests that changing the head position during an interaction should send intuitive signals and could be integrated within expressive systems. Such signals could be used by a human to assess whether an interaction is successful.

The results obtained in this study are consistent with existing studies looking at body postures displayed by humans [Kleinsmith, Bianchi-Berthouze and Steed 2011; Kleinsmith, De Silva and Bianchi-Berthouze 2006]. Taken together, this reinforces the results of the first study as it suggests that embodiment has no effect on the identification of emotional body language. If this is the case, an unrealistic humanoid robot such as Nao should also be able to express richer and more complex emotions than the one tested in this study.

9. AUTOMATIC GENERATION OF EMOTIONAL EXPRESSIONS

Using the results of study 2, an 'Affect Space' based on the circumplex model of emotion [Russell 1980] was defined for body language. According to Russell's circumplex model of emotions, emotional experiences depend on two major dimensions, Arousal and Valence. The postures were selected by looking at how head position affected the interpretation of the emotion displayed. In order to generate the two-dimensional space, four emotions were chosen from this study. Happiness was chosen as it was the positive emotion conveying the highest level of Arousal. Pride was chosen because it was the positive emotion conveying the lowest level of Arousal. For the negative emotions, Fear was chosen as it was conveying the highest level of Arousal. Sadness was chosen as it was conveying the lowest level of Arousal. A neutral and stable pose was developed and added to the set.

Finally, the axes of the Affect Space were built by placing in opposition the most positive and aroused posture with the most negative and non-aroused key pose. Similarly, the most negative and aroused key pose was placed in opposition with the most positive and non-aroused key pose. New postures are generated by calculating the weighted mean of the joint angles from up to three postures taken from the set. The four key poses were slightly modified to improve the stability of the robot and to ensure it would not fall on account of a bad combination. The resulting system creates postures along the two dimensions of the circumflex model (Arousal and Valence) and insures that all the blending was consistent with the effect found when manipulating the position of the head. This approach is interesting as it produces a wide range of different emotional expressions easily and quickly. These animations are fully configurable and use only a small amount of memory. Each key frame is computed "on the fly". Another benefit is the ability to change from one emotion to another without the need to display an intermediary neutral pose. It is also possible to easily add new basic key poses for every emotion, producing a very wide range of emotional expressions.

However, these types of expressions need to be tested as their interpretation may differ from the intended one. Although 'blending' works for facial expression, it is not evident that 'blending' two negative body expressions would result in a third different negative expression. Therefore, an experiment investigating how such key poses are perceived was conducted to validate the affect space.

10. VALIDATING THE GENERATED EXPRESSIONS

The experiment was designed to test how key poses generated by the Affect Space are interpreted and whether these interpretations are consistent with the postures' position in the model. The experiment used a within-subjects design with one dependent variable (Emotion Displayed).

Four dependent variables were defined to explore the Affect Space: Primary Emotion, Secondary Emotion, Arousal and Valence (See section 1 for definition). Primary Emotion and Secondary Emotion were used to test whether it was possible for participants to interpret the key pose displayed. Arousal and Valence were used to investigate the position of each tested key pose' in the Affect Space.

The main question tested was:
Is the interpretation of the key poses displayed consistent with their positions in the Affect Space?

10.1. Participants

23 Participants were recruited, mostly students from the University of Portsmouth (7 females and 16 males) ranging in age from 19 to 49 (M=27.22, SD=7.80). Participants were entered in a raffle to win an iPhone in exchange for participation.

10.2. Material and Apparatus

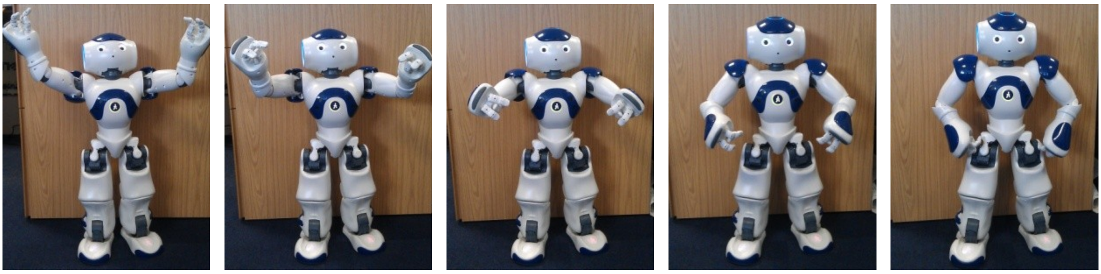

Fig. 12. Five Key poses generated by the system. From 100% Happiness (top left) to 100% Pride (top right). Middle pictures are combination of the two.

The selected key poses from Study 2 (Section 9) were modified to improve the stability of the robot and to ensure that it would not fall on account of a bad combination. Sixteen additional poses were then generated using the affect space. Each emotion was 'blended' with its 'neighbours' at three different levels (100%, 70%/30%, 50%/50%) (Figure 12). To limit the number of key poses being assessed by each participant, the neutral position was blended with all emotions at 50%/50% only.

10.3. Procedure

All participants were tested by the same experimenter in individual sessions. Each session began by obtaining consent, then participants watched and assessed the 20 poses. Each pose was displayed only once in a randomized order. For each pose, participants were asked to make an 'open' interpretation. They had to categorize it by choosing one emotion among Happiness, Pride, Excitement, Fear, Anger, Sadness and Neutral. Eventually they had the possibility to add a secondary emotion. Participants also rated Valence (Negative/Positive) and Arousal (Low Energy/High Energy) on a 10-point Lickert scale. Once all the poses had been assessed, participants were fully debriefed regarding the purpose of the study. The whole session took around 30 minutes.

11. SUMMARY OF RESULTS AND DISCUSSIONS

In this section, only a summary of the results are presented. A more complete analysis can be found in [Beck et al. 2010].

11.1. Identification of the Basic Set

Table IV. Recognition rate of the set of posture used (Chance level would be 14%)

Sadness	Fear	Pride	Happiness	Neutral
74%	87%	96%	44%	78%

Since the original postures were slightly modified and a neutral posture was added, it was necessary to check whether it was still possible for participants to correctly identify them.
Table IV confirms that participants were able to interpret all the postures used in the set. As in the previous study, Happiness was most commonly misinterpreted as Excitement (by 26% of participants). In the context of this experiment this was a positive result as it confirmed that the key pose showing Happiness was likely to be perceived as positive and aroused.

11.2. Interpretations of the Generated Key Poses

The interpretations of the key poses displayed were consistent with their positions in the model. The negative key poses, that were automatically generated, were interpreted as negative whereas the positive ones were interpreted as positive. Moreover, for most of the key poses, the primary interpretation was consistent with the 'blend' of emotions being displayed.

The recognition rates confirmed that it is possible to interpret these emotions when displayed by a humanoid robot and that the lack of facial expression is not a barrier to expressing emotions. Moreover, the results show that it was possible for participants to successfully recognize the key poses generated by the system. For instance, the key poses created by blending 70%/30% of different emotions were interpreted in a manner consistent with the primary emotions being displayed. This suggests that it is possible to create variations of an emotional expression using the Affect Space while maintaining the way it is perceived. In other words, this method can be used to automatically generate different expressions for an emotion.

However, the results suggest that the key poses created by blending emotion at 50%/50% were more difficult to interpret. For instance 50% Happiness 50% Pride was interpreted by 30% of the participants as Neutral and by another 30% as Happiness. However, looking at the value of Valence and Arousal, the position of the key pose was still consistent with the model. This was further suggested by the answers to the open question. For instance, four participants described the key pose as 'Welcoming', 'Embracing' or 'Wants a hug'. This shows that it was perceived as predicted (Positive but less aroused than 100% Happiness). Similarly, 50% Fear 50% Sadness was interpreted as Fear by only 35% of the participants. Looking at the value of Valence and Arousal, the key pose's position was still consistent with the model.

The interpretation of the key poses thus suggests that the Affect Space created can be used to enrich the expressiveness of the robot. It could also be used to avoid always displaying the exact same expression for an emotion while still being understandable.

11.3. Perceived Valence

In order to investigate how the blended postures were perceived, the different key poses were compared in pairs using Two Ways Repeated Measures ANOVAs. Assumptions of normality, homogeneity of variance and sphericity were met.

As expected, Happiness was perceived as significantly more positive (*Valence*) than Sadness ($F(1,22)=69.51$, $p<0.01$, Partial $\eta^2=0.76$) or Fear ($F(1,22)=73.59$, $p<0.01$, Partial $\eta^2=0.77$). Pride was perceived as more positive than Sadness ($F(1,22)=106.55$, $p<0.01$, Partial $\eta^2=0.83$) or Fear ($F(1,22)=164.14$, $p<0.01$, Partial $\eta^2=0.88$). There was no difference between Happiness and Pride ($F(1,22)=0.04$, $p=0.84$, $\eta^2=0.00$). Similarly, there was no difference between Sadness and Fear ($F(1,22)=0.68$, $p=0.42$, Partial $\eta^2=0.03$).

The results of the comparisons made between the different postures shows that overall the perceived *Valence* of the Positive Aroused area (i.e. Happiness) was not affected by the changes in postures, neither by adding elements of Fear nor Pride. However, the 'Positive non-aroused' (i.e. Pride) area was affected by the changes when adding elements of Happiness.
For the Negative Aroused (i.e. Fear) and the Negative Non Aroused (i.e. Sadness) areas, the results show that the key poses from the basic set were perceived as the most negative (*Valence*) ones.

The algorithm did not create 'aberrant' postures. The perceived *Valence* was always consistent with the emotions being displayed. This was also confirmed by the interpretations of the emotions, which were consistent with the intended display.

However, there were some unexpected results regarding the perceived Valence of the negative key poses generated. The results show that the key poses generated by blending Fear and Sadness were perceived as less negative than the original ones. The model predicted no change in *Valence*. The key pose 100% Fear and the key pose 100% Sadness may have been perceived as extreme occurrence, prototypical displays, of these emotions. This would explain why they were perceived as more negative than the generated ones, which are not prototypical. Nevertheless, the generated key poses were still interpreted as negative. The organization of the Affect Space will be modified to take this into account. Moreover, the Affect Space was tested with key poses and it is expected that adding movements will further improve the expressivity of the system.

11.4. Perceived Arousal

As expected, Happiness was perceived as significantly more *Aroused* than Pride ($F(1,22)=10.27$, $p<0.01$, Partial $\eta^2=0.32$) or Sadness ($F(1,22)=166.84$, $p<0.01$, Partial $\eta^2=0.88$). Fear was perceived as more *Aroused* than Sadness ($F(1,22)=47.13$, $p<0.01$, Partial $\eta^2=0.68$). Moreover, Pride was perceived as more Aroused than Sadness ($F(1,22)=30.55$, $p<0.01$, Partial $\eta^2=0.58$). Happiness was perceived as more Aroused than Fear ($F(1,22)=5.46$, $p<0.05$, Partial $\eta^2=0.20$).

As in the previous section, the results show that perceived *Arousal* is consistent with the prediction of the model. In other words, 'blending' an aroused emotion with a non-aroused one either decreases the perceived Arousal or does not affect it. Similarly, blending a non-aroused emotion with an aroused one increases the perceived *Arousal* or does not affect it. Moreover, for each emotion, there was a decrease (significant or a trend) in Arousal when it was blended with the neutral key pose.

The generated key poses were consistent with the predictions made by the Affect Space. It is possible to increase or decrease the perceived Arousal by adding elements of an aroused or un-aroused posture. For instance, the key pose 50% Fear 50% Sadness, was interpreted as Neutral. It was however rated as more aroused than Sadness and less aroused than Fear.

However, the results also show that the anticipated position of the postures needs to be corrected. For instance, 100% Pride was completely misplaced, as it conveyed a higher level of Arousal. Because of this, the Affect Space generated for this study did not cover the 'positive non-aroused area'. It will be necessary to complete it with a non-aroused positive posture.

So far, only key poses have been tested and Arousal is known to be related to the speed of movements [Saerbeck and Bartneck 2010]. It is therefore expected that the model will benefit from incorporating motion varying in speed depending on the robot's arousal.

12. CONCLUSION

Consistent with the argumentation developed by Cañamero [2006], the Uncanny Valley model was found to be incomplete. It was found that individual differences affect the perceived believability of a character and that a simple one-to-one relationship between believability and realism does not correspond to the reality. More precisely emotional intelligence and previous video game experience were found to be related with believability. It would be interesting to investigate if other factors have an effect 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 individual differences and how they interact with the perception of realistic characters.

The empirical studies 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. This is relevant to robots as it suggests that even simplified humanoid robots can use body language to express emotions and that a non realistic humanoid robot should use stylized displays of emotions.

This was further confirmed by the second empirical study, which demonstrates that expressive key poses can be accurately interpreted and that, as for humans, head position is highly expressive. Emotional body language is an appropriate medium for expressing emotions even with simplified bodies such as the one of a humanoid robot. Moreover, the study of emotional body language can benefit from the use of humanoid robots such as Nao as they can display rich emotional body language, similar to the one displayed by humans, without being confounded by other modalities such as facial expressions.

The third study shows that it is possible to interpret expressions generated by blending key poses along a continuous model of emotions. This suggests that the approach can be used to enrich, at a low cost, the expressiveness of humanoid robots. However, the exact position in the Affect Space of the generated expressions still needs to be clearly assessed. Moreover, this method can generate short emotional animation 'on the fly'. However, the interpretation of such animations remains to be tested. The final version will consider acceleration and curvature, as it has been established that these parameters are related to arousal and valence [Saerbeck and Bartneck 2010]. These additions should improve the expressiveness of robots.

Moreover, since the overall purpose of communicating the emotional state of the robot is to facilitate interaction with humans, the effectiveness of the Affective Space will be assessed in the context of real-time interactions. The evaluation will consider the recognition of the postures being displayed as well as their effect on the interaction. It is expected that the widening of the range of emotional expressions of the robot, will help human partners interact with the robot intuitively.

AKNOWLEDGMENTS

Part of this work was conducted while Aryel Beck was at the School of Creative Technologies, University of Portsmouth.

The authors would like to thank the director Romain Beck and the actor Will Pinchin for their excellent work.

This work is partly funded by the EU FP6 Feelix Growing project (grant number IST-045169), and partly by the EU FP7 ALIZ-E project (grant number 248116).

REFERENCES

ALDEBARAN 2010. http://www.aldebaran-robotics.com/.
AMS 2011. http://www.actorsmovementstudio.com/rasaboxes/.
ARGYLE, M. 1988. *Bodily communication*. Routledge.
ATKINSON, A.P., DITTRICH, W.H., GEMMEL, A.J. AND YOUNG, A.W. 2004. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception 33*, 717-746.
AYLETT, R., PAIVA, A., WOODS, S., HALL, L. AND ZOLL, C. 2005. Expressive Characters in Anti-Bullying Education. *publication13280*.
BATES, J. 1994. The role of emotion in believable agents. *Communications of the ACM 37*, 122-125.
BECK, A., CAÑAMERO, L. AND BARD, K. 2010. Towards an affect space for robots to display emotional body language. In *Proceedings of the Ro-Man*, Viareggio, Italy2010 IEEE.
BECK, A., HIOLLE, A., MAZEL, A. AND CANAMERO, L. 2010. Interpretation of Emotional Body Language Displayed by Robots. In *Affine 2010* ACM, Firenze, Italy.
BECK, A., STEVENS, B. AND 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. AND BARD, K. 2009. Comparing perception of affective body movements displayed by actors and animated characters. In *AISB*, N. BERTHOUZE, M. GILLIES AND A. AYESH Eds., Edinburgh, UK.
BREAZEAL, C. 2002. *Designing sociable robots*. MIT press.
BRENTON, H., GILLIES, M., BALLIN, D. AND CHATTING, D. 2005. The Uncanny Valley: does it exist and is it related to presence ? *Presence-connect http://www.presence-research.org/*.
CANAMERO, L. AND FREDSLUND, J. 2001. I show you how I like you - Can you read it in my face? *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans 31*, 454-459.
CASTELLANO, G. 2008. Automated analysis of body movement in emotionally expressive piano performances. *Music Perception 26*, 103-119.
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*, 355-359.
CHAPMAN, N. AND CHAPMAN, J. 2004. *Digital multimedia*. John Wiley & Sons, Ltd.
COULSON, M. 2004. Attributing Emotion To Static Body Postures: Recognition Accuracy, Confusions, And Viewpoint Dependence. *Journal of nonverbal behavior 28*, 117-139.
DANTZER, R. 2005. *Les Emotions*. Presses universitaires de France, Paris.
DAUTENHAHN, K. AND HURFORD, J. 2006. Interaction studies *7*.
DAVIDSON, J.W. 1994. What type of information is conveyed in the body movements of solo musician performers? *Journal of Human Movement Studies 6*, 279-301.
DE GELDER, B. 2009. Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the the Royal Society B: Biological Sciences 364*, 3475-3484.
DE MELO, C. AND PAIVA, A. 2008. Expression of emotion is virtual humans using lights, shadows, composition and filters. In *Proceedings of the IVA 2008*, Tokyo, Japan2008, H. PRENDINGER, J. LESTER AND M. ISHIZUKA Eds.
DE SILVA, R.P. AND BIANCHI-BERTHOUZE, N. 2004. Modeling human affective postures: an information theoretic characterization of posture features. *Computer Animation and Virtual Worlds 15*, 269-276.
DEN STOCK, J.V., RIGHART, R. AND DE GELDER, B. 2007. Body expressions influence recognition of emotions in the face and voice. *Emotion 7*, 487-494.
EKMAN, P., FRIESEN, W.V. AND HAGER, J.C. 2002. *Facial Action Coding System, the manual*. A Human Face, Salt Lake City.
ENZ, S., ZOLL, C., VANNINI, N., WATSON, S., AYLETT, R., HALL, L., PAIVA, A., WOLKE, D., DAUTENHAHN, K., ANDRE, E. AND RIZZO, P. 2008. Virtual role-play in the classroom - Experiences with FearNot! In *Proceedings of the eChallenges*, Stockholm, Sweden2008.
GILLIES, M., PAN, X., SLATER, M. AND SHAWE-TAYLOR, J. 2008. Responsive listening behavior. *Computer animation and virtual worlds 19*, 579-589.
GRAMMER, K., BERNHARD, F., ELISABETH, O., MICHAELA, A., INES, B. AND PHILIPP, M. 2004. The representation of self reported affect in body posture and body posture simulation. *Collegium Antropologicum 28*, 159-173.
HANSON, D., OLNEY, A., PEREIRA, I.A. AND ZIELKE, M. 2005. Upending the uncanny valley. In *Proceedings of the American association for artificial intelligence (AAII) conference*, Pittsburgh, PA, USA2005 AAII.

HIOLLE, A., CAÑAMERO, L., ANDRY, P., BLANCHARD, A. AND GAUSSIER, P. 2010. Using the interaction rhythm as a natural reinforcement signal for social robots: A matter of belief. In *International Conference on Social Robotics*, SPRINGER Ed., Singapour, 81-89.
JAMY, L., MARK, C., SACHI, M. AND MICHIAKI, Y. 2009. Emotions and messages in simple robot gestures. In *Human-Computer Interaction. Novel Interaction Methods and Techniques* Springer, Berlin/Heidelberg, 331-340.
KLEINSMITH, A., BIANCHI-BERTHOUZE, N. AND STEED, A. 2011. Automatic Recognition of Non-Acted Affective Postures. *IEEE Trans Syst Man Cybern B Cybern*.
KLEINSMITH, A., DE SILVA, P.R. AND BIANCHI-BERTHOUZE, N. 2006. Cross-cultural differences in recognizing affect from body posture. *Interacting with Computers 18*, 1371-1389.
KRUSKAL, J.B. AND WISH, M. 1978. *Multidimensional scaling*. Sage publications, Beverly Hills and London.
MACDORMAN, K.F., GREEN, R.D., HO, C.C. AND KOCH, C. 2009. Too real for comfort: Uncanny responses to computer generated faces. *Computers in human behavior 25*, 695-710.
MCDONNELL, R., JORG, S., MCHUGH, J., NEWELL, F. AND O'SULLIVAN, C. 2008. Evaluating the emotional content of human motions on real and virtual characters. In *5th symposium on Applied perception in graphic and visualization*, S. CREEM-REGEHR AND K. MYSZKOWSKI Eds. ACM, Los Angeles, California, 67-74.
MORI, M. 1970. Bukimi no tani [The un-canny valley] . *Energy 7*, 33-35.
NASS, C. AND MOON, Y. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues 56*, 81-103.
PELACHAUD, C. 2009. Modelling multimodal expression of emotion in a virtual agent. *Philos Trans R Soc Lond B Biol Sci 364*, 3539-3548.
ROETHER, C., OMLOR, L., CHRISTENSEN, A. AND GIESE, M. 2009. Critical features for the perception of emotion from gait. *Journal of Vision*, -.
RUSSELL, J.A. 1980. A circumflex model of affect. *Journal of Personality and Social psychology 39*, 1161-1178.
SAERBECK, M. AND BARTNECK, C. 2010. Perception of affect elicited by robot motion. In *Proceedings of the Human-Robot Interaction (HRI2010)*, Osaka2010, ACM/IEE Ed. ACM/IEE, 53-60.
SAKAGUCHI, H. AND SAKAKIBARA, M. 2001. Final Fantasy: The Spirits Within Sony Pictures, Japan.
SCHERER, K.R. 2005. What are emotions? And how can they me measured? *Social Science Information 44*, 693-727.
SCHOUWSTRA, S. AND HOOGSTRATEN, J. 1995. Head position and spinal position as determinants of perceived emotional state. *Perceptual and Motor Skills 81*, 673-674.
SCHRÖDER, M. 2001. Emotional speech synthesis - A Review. In *Proceedings of the Eurospeech 2001*, Aalborg2001, 561-564.
SCHUTTE, N.S., MALOUFF, J.M., HALL, L.E., HAGGERTY, D.J., COOPER, J.T., GOLDEN, C.J. AND DORNHEIM, L. 1998. Development and validation of a measure of emotional intelligence. *Personality and Individual Differences 25*, 167-177.
THOMAS, F. AND JOHNSTON, O. 1995. *The illusion of life*. Abbeville Press, New-York.
WALLBOTT, H.G. 1998. Bodily expression of emotion. *European Journal of Social Psychology 28*, 879-896.
ZECCA, M., MIZOGUCHI, Y., ENDO, K., IIDA, F., KAWABATA, Y., ENDO, N., ITOH, K. AND TAKANISHI, A. 2009. Whole body emotion expressions for KOBIAN humanoid robot ; preliminary experiments with different emotional patterns. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, 381-386.
ZEMECKIS, R. 2004. The Polar Express Warner Bross, USA.