Children Interpretation of Emotional Body Language Displayed by a Robot

Aryel Beck[1], Lola Cañamero[1], Luisa Damiano[1], Giacomo Sommavilla[2], Fabio Tesser[2] and Piero Cosi[2]

[1]Adaptive Systems Research Group, School of Computer Science & STRI, University of Hertfordshire, United Kingdom
{a.beck2, l.canamero, l.damiano}@herts.ac.uk
[2]Institute of Cognitive Sciences and Technologies, Padova, Italy
{giacomo.sommavilla, fabio.tesser, piero.cosi}@pd.istc.cnr.it

Abstract. Previous results show that adults are able to interpret different key poses displayed by the robot and also that changing the head position affects the expressiveness of the key poses in a consistent way. Moving the head down leads to decreased arousal (the level of energy), valence (positive or negative) and stance (approaching or avoiding) whereas moving the head up produces an increase along these dimensions [1]. Hence, changing the head position during an interaction should send intuitive signals which could be used during an interaction. The ALIZ-E target group are children between the age of 8 and 11. Existing results suggest that they would be able to interpret human emotional body language [2, 3].
Based on these results, an experiment was conducted to test whether the results of [1] can be applied to children. If yes body postures and head position could be used to convey emotions during an interaction.

1 Introduction

The ALIZ-E project aims to contribute to the development of integrated cognitive systems capable of naturally interacting with young people in real-world situations, with a specific goal of supporting children engaged in a residential diabetes-management course. Fundamental to making human-robot interaction natural and integrated into the fabric of our lives, is that the robot can establish itself cognitively in the long term. Only if interaction provides a sense of continuity over longer periods of time, can it provide the resonance necessary for a constructive relationship between human and robot. It is commonly acknowledged that learning, adaptation, emotion, multi-modal dyadic and group interactions will be necessary to achieve this goal, but the field has not yet been presented with conclusive design paradigms, algorithms and test results showing how a robot can enter and successfully maintain an interaction spread beyond the current single episode interaction frame and stretching over several days. The work reported in this paper focuses on the emotional aspect. More precisely, it is concerned with developing methods that will enable a robot to display emotions in a way that can be readily interpreted by children during an interaction.

Existing work in achieving expressive agents is difficult to apply to humanoid robots such as Nao. Work has been conducted on computer agents that achieve responsive behaviors using bodily expressions. For instance, Gillies et al. (2008) have designed a method to create responsive virtual humans that can generate expressive body language while listening to a human. Their expressions are based on motion capture data [4]. However, it would be difficult and tedious to transfer this method onto robots directly as they cannot reproduce the movements recorded by motion capture as smoothly as virtual humans or without causing the robot to often lose its balance. Expressive robots have also been successfully created. For instance, Kismet expresses emotions through its face [5]. Its expressions are based on nine prototypical facial expressions that 'blend' (interpolate) together along three axes: Arousal, Valence and Stance. Arousal is defined as the level of energy. Valence specifies how positive or negative the stimulus is. Stance reflects how approachable the stimulus is. This method defines an Affect Space, in which expressive behaviours span continuously across these three dimensions, allowing a wide range of expressions [5]. The method is interesting for its simplicity; however, the stance dimension may be problematic as it is not related to any accepted model of emotions. Modeling expressive emotions based on notion that are not validated in psychology may be problematic for long term interaction outside the laboratory. It may result in displaying inappropriate behaviour which could be counterproductive for the interaction. Moreover, for many robots such as Nao the same Affect Space cannot be directly applied as they do not have the ability to display facial expressions. The only medium available for such robots to express emotions is their bodies and voices.

2 Body Language as a Modality for Robot to display Emotions

It has been shown that body language can be interpreted accurately without facial or vocal cues [6-8]. These results suggest that a humanoid robot, such as Nao, should be able to display emotions using its body. This is further suggested by traditional animation which focuses on the display of emotion through the body in order to increase believability. It has been codified as a rule in classical animations: "the expression must be captured throughout the whole body as well as in the face" [9]. Theatre follows a similar principle, by asking actors to become, in Artaud's words, "athletes of the emotions". Moreover, a large part of an actor's training addresses the non-verbal expression of emotions. These suggest that emotions such as fear, anger, happiness, stress, etc., could be readable when expressed through the body of Nao. However, most of the research on the psychology of emotions has focused on facial expressions. This makes it difficult to create an expressive systems based on bodily expressions. There is no equivalent to the Facial Action Coding System [10] for body expressions. Researchers have categorized the different types of body language, depending on how they occur. The categorization presented below was created from [11, 12] and classifies body language into three different areas.

Postures: Postures are specific positioning that the body takes during a timeframe. It has been established that postures are an effective medium to express emotion. For instance, De Silva et al. (2004) investigated cross-cultural recognition of four emotions (anger, fear, happiness, sadness) through interpretations of body postures. They built a set using actors to perform emotional postures and showed that it was possible for participants to correctly identify the different emotions [13]. Thus, a humanoid robot displaying emotion should take up postures appropriate to the emotion.

Movement: It has been shown that many emotions are differentiated by characteristic body movements, and that these are effective cues for judging the emotional state of other people even in the absence of facial and vocal cues [14]. Thus, a Nao robot displaying emotion should also do so during, and via, motion. Body movements include the movements themselves as well as the manner in which they are performed (i.e. movement speed, dynamics, curvature, etc.). The movements' dynamics have been shown to contribute to the emotional expression. For instance, Wallbott (1998) compared body language displayed by actors portraying different emotional states and found significant differences in the movement dynamics as well as in the type of movements performed across emotions [15]. Pollick et al (2001) investigated affect from point-light display of arm movements and found that activation is a formless cue that relates directly to the kinematics of the movements [16]. These studies are interesting because they show that dynamics is an essential component of an emotional expression.

Proxemics: It is the distance between individual during a social interaction. It is also indicative of emotional state. For example, angry people have a tendency to reduce the distance during social interaction. The problem is that this would also be the case between intimate people. Hence, proxemics cannot therefore be considered as an emotional expression in itself but is required to complete a representation of emotional behaviour and could be an interesting addition for the expressivity of a robot.

3 Previous Results

In animation, one of the established methods for creating convincing and believable displays consists in starting from the creation of expressive key poses (i.e. postures) rather than body language in motion [9]. 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 by changing the position of a subset of joints, they can then be used to drive the expressive behaviours of the robot.

Previous work focused on validating a set of key poses and on testing the effect of moving the head up or down in a range of different key poses [1]. The position of the head was chosen because of its importance regarding the expression of emotions [17, 18]. Moreover, animation emphasizes the importance of creating strong silhouette [18, 19] and it is expected that manipulating the head position will considerably change a robot's silhouette.

This experiment showed that it was possible for adults to interpret the different key poses displayed by the robot and also that changing the head position affects the expressiveness of the key poses in a consistent way. 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 [1]. This suggests that changing the head position during an interaction should send intuitive signals which will be used, for example, to indicate whether an interaction is successful.

These results were established with adults. However, the ALIZ-E project focuses on children and it is therefore necessary to test whether they can be extended to such a specific population. The results could depend on cultural and age differences.

4 Research Question

According to Boone and Cunningham's research on developmental acquisition of emotion decoding from expressive body movement [2, 3], as children begin to produce certain actions, they have access to the perceptual expressive cues associated to these actions. In turn, this can lead to effective cue utilisation. Boone and

Cunningham experiment shows that, with respect to adults, it is possible to associate cues in naturally generated dance expression to specific emotions, and that children, from 8 years of age, can recognise them for the target emotions of happiness, sadness, anger, and fear [3]. However, existing studies have also shown that emotional recognition continues to develop during the adolescence [20]. Additionally, research in the perception of robots, suggests that there may be differences in the way children and adults perceive them [21]. It is therefore not evident that children and adults would interpret the body language displayed by a robot similarly. Thus, the purpose of the study reported in this paper was to test the results of [1] with children and to investigate whether the head position could be used to convey different emotions to such a specific population.

5 The Experiment

The experiment setting was defined to be as similar as possible to the one used with adult participants [1]. It 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 (Table 1).

Head position had three levels (Up, Down, and Straight), defined as the head position relative to the chest. One dependent variable was defined to explore the Affect Space: *Correct Identification. It* was used to test whether or not it was possible for participants to interpret the emotion of the key poses. Although the study conducted on adults was investigating *Arousal*, *Valence* and *Stance* as well, it was decided to remove them from this study because of the age difference.

The three main questions tested were:
(Q1) Are children as accurate as adults in identifying the key poses displayed by Nao?
(Q2) What is the effect of changing the head position on the interpretation and perceived place of a key pose in the Affect Space?
(Q3) 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?

5.1 Participants

24 Children (13 females, 11 males) were recruited from the school "scuola media Dante Alighieri" (Italy) ranging in age from 11 to 13 (M=12, SD=0.3).

5.2 Material

The same material as in the study conducted with adults was reused. The reader can refer to [1] for a detailed report on the construction of the material. The platform chosen for this experiment was Nao, a humanoid robot with 25 degrees of freedom.

The experimental poses were generated by systematically altering the head positions of 6 emotional key poses. 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 Positions*).

5.3 Procedure

The same experimenters tested all participants in groups of four. Participants were given full explanation regarding the questionnaire that they were expected to answer and were instructed to *"imagine that the robot is reacting to something"*.
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 group of participants. 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. When all the poses were assessed, participants were fully debriefed. The sessions lasted approximately 30 minutes.

6 Results

6.1 Identification of the emotion

Table 1. Percentage of participants who correctly identified the emotional key pose at least once (Chance level would be 42%)

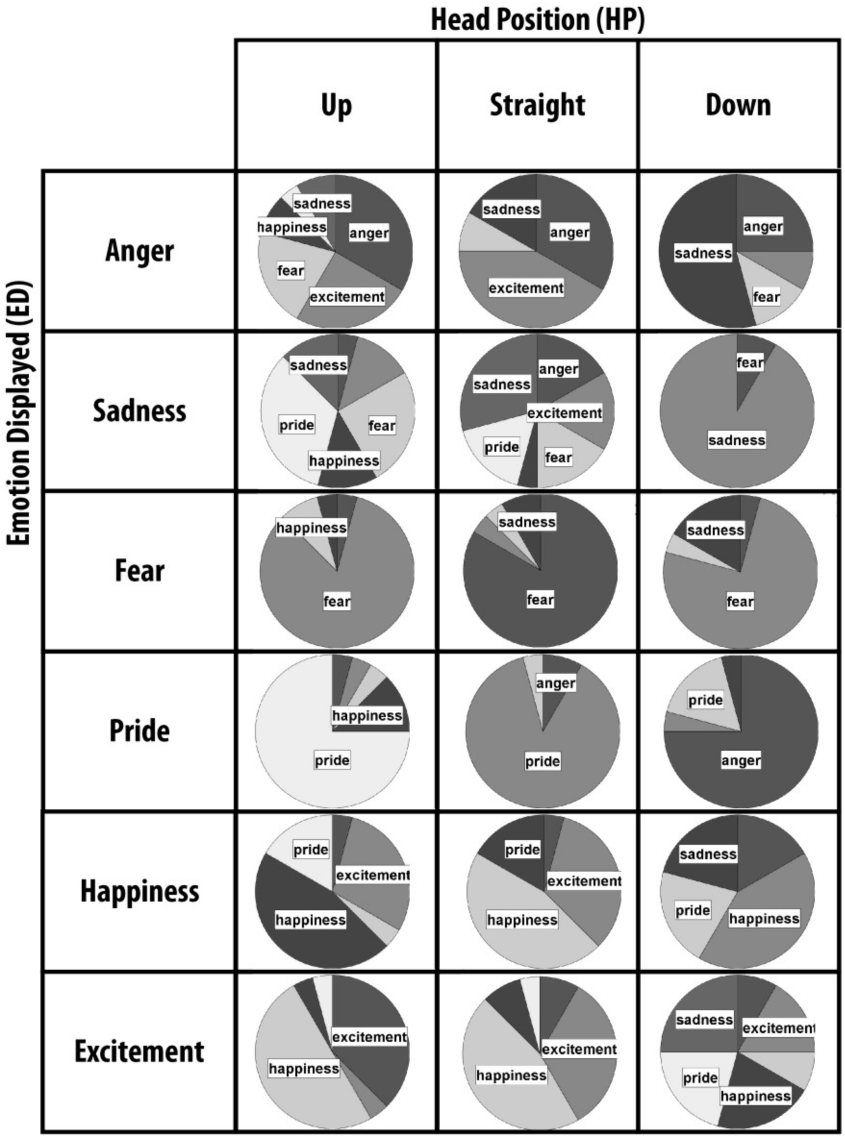
Pride	Happiness	Excitement	Fear	Sadness	Anger
100%	83%	63%	92%	92%	58%

Repeated Measures ANOVA (6 *Emotion Displayed* x 3 *Head Position)* was conducted on *Correct Identification*. *Emotion Displayed* had a significant effect on *Correct Identification (CI)* ($F(5,115)=12.03$, $p<0.01$, Partial $\eta^2=0.34$). *Head Position* had no significant main effect on *Correct Identification* ($F(2,46)=1.45$, $p=0.25$, Partial $\eta^2=0.06$).

These results indicate that participants' performance was different across emotions. Table 1 shows that the children correctly identified each emotion from viewing only the key pose of Nao. Recognition rates were above chance level although they varied from 58% for anger, to 100% for pride. Chance level would be $(1-(5/6)^3)*100= 42\%$.

6.2 Effect of head position on the interpretation

Table 2. Effect of Head Position on the Interpretations of the body language displayed.



There was a significant interaction between *Emotion Displayed* and *Head Position* ($F(10,230)=9.32$, $p<0.01$, Partial $\eta^2=0.29$). This indicates that the effect of *Head Position* on *Correct Identification* depended on the individual emotion being displayed. Therefore, the effect of *Head Position* were considered separately for each emotion and are reported in Table 2. It shows how the emotional interpretations of the displays shifted as a function of both the *Emotion Displayed* and the *Head Position*. The patterns found in this study are comparable to the one found with adult participants [1]. Participants were better at interpreting the negative emotions when the *Head Position* was Straight or Down. Participants were better at interpreting the positive emotions when the *Head Position* was Up.

7 Discussions

The first goal of the study was to test the expressivity of the key poses displayed by the robot with children. As with adults, the results show that the children who participated in the study were far better than chance level at interpreting the different key poses taken by the robot (Table 1). 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 2). These results clearly show that it is possible for children to interpret emotions displayed by a humanoid robot and that the lack of facial expression is not a barrier to expressing emotions. This 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.

As in [1], Head Position had a strong effect on the interpretation of the key poses being displayed (Table 2). For instance, children's interpretations of the Pride display were very similar to those of the adults. More precisely, it was interpreted as Pride when the head was up or straight. However, with the head down, a majority of children interpreted it as anger (Table 2). Fear was not affected by the change in Head Position and was correctly interpreted in all conditions both by the adults and the children. This further suggests that the interpretation of the key poses was similar in the adult and children's testing conditions.

Interestingly, children were less accurate than adults at interpreting the 'anger' key pose (58% vs. 89%). This difference could be due to cultural or age differences or to the different settings between the two experiments. This is an interesting issue and should be explored in future research as it is not possible to draw definitive conclusions from this study. Moreover, it is important to highlight that the material used for this study is prototypical and was intentionally selected to be expressive. This is appropriate within the ALIZ-E project; however, it is likely that the use of prototypical expressions had an effect on the results and on the similarities of the interpretations that were found in this study.

8 Conclusion

As with adults, it was found that moving the head up increased the identification of some emotions (pride, happiness, and excitement), whereas moving the head down increased correct identification for other displays (anger, sadness). Fear, however, was well identified regardless of Head Position.

This has design implication for improving emotional body language displayed by robots. The results of this study suggest that the expressivity of the negative emotions (anger and sadness) can be improved by moving the head down while the expressivity of the positive emotion (happiness, excitement and pride) can be improved by moving the head up. These results have already been successfully integrated in an automated expressive system [22]. The robot can automatically change its head position to express changes in its internal state.

Future work will explore the effect of moving the different parts of the body on the interpretation of the body language displayed as well as adding dynamic elements to the expressions. If similar results can be established for the other parts of the body, it will be possible to create a rich Affect Space for humanoid robots.

Acknowledgments. The authors would like to thank the school “scuola media Dante Alighieri” for hosting the experiment.
This work is funded by the EU FP7 ALIZ-E project (grant number 248116).

References

1. Beck, A., L. Cañamero, and K. Bard, *Towards an affect space for robots to display emotional body language*, in *Ro-Man*2010, IEEE: Viareggio, Italy.
2. Boone, R.T. and J.G. Cunningham, *The Attribution of Emotion to Expressive Body Movements: A Structural Cue Analysis*, 1996: Manuscript.
3. Boone, R.T. and J.G. Cunningham, *Children's decoding of emotion in expressive body movement: the development of cue attunement.* Dev Psychol, 1998. **34**(5): p. 1007-16.
4. Gillies, M., et al., *Responsive listening behavior.* Computer animation and virtual worlds, 2008. **19**(5): p. 579-589.
5. Breazeal, C., *Designing sociable robots*. Intelligent Robotics & Autonomous Agents2002: MIT press.
6. Beck, A., B. Stevens, and K. Bard. *Comparing perception of affective body movements displayed by actors and animated characters*. in *AISB*. 2009. Edinburgh, UK.

7. de Gelder, B., *Towards the neurobiology of emotional body language.* Nature reviews Neuro-Science, 2006. **7**(3): p. 242-249.
8. Kleinsmith, A., P.R. De Silva, and N. Bianchi-Berthouze, *Cross-cultural differences in recognizing affect from body posture.* Interacting with Computers, 2006. **18**(6): p. 1371-1389.
9. Thomas, F. and O. Johnston, *The illusion of life*1995, New-York: Abbeville Press.
10. Ekman, P., W.V. Friesen, and J.C. Hager, *Facial action coding system. The manual*2002, Salt Lake: Human Face.
11. Cassell, J., *Nudge nudge wink wink: elements of face-to-face conversation*, in *Embodied Conversational Agents*, J. Cassell, et al., Editors. 2000, MIT Press: Cambridge. p. 1-27.
12. Vinayagamoorthy, V., et al. *Building Expression into Virtual Characters*. in *Eurographics 2006*. 2006. Vienna: Proc. Eurographics.
13. De Silva, P.R. and N. Bianchi-Berthouze, *Modeling human affective postures: an information theoretic characterization of posture features.* Computer Animation and Virtual Worlds, 2004. **15**(3-4): p. 269-276.
14. Atkinson, A.P., et al., *Emotion perception from dynamic and static body expressions in point-light and full-light displays.* Perception, 2004. **33**: p. 717-746.
15. Wallbott, H.G., *Bodily expression of emotion.* European Journal of Social Psychology, 1998. **28**(6): p. 879-896.
16. Pollick, F.E., et al., *Perceiving affect from arm movement.* Cognition, 2001. **82**(2): p. 51-61.
17. Kleinsmith, A., N. Bianchi-Berthouze, and A. Steed, *Automatic Recognition of Non-Acted Affective Postures.* IEEE Trans Syst Man Cybern B Cybern, 2011.
18. Schouwstra, S. and J. Hoogstraten, *Head position and spinal position as determinants of perceived emotional state.* Perceptual and Motor Skills, 1995. **81**(2): p. 673-674.
19. Maestri, G., *Digital character animation 3*, ed. E. Kissane and K. Kalning2006, Berkeley: New Riders.
20. Tonks, J., et al., *Assessing emotion recognition in 9-15-years olds: preliminary analysis of abilities in reading emotion from faces, voices and eyes.* Brain Inj, 2007. **21**(6): p. 623-9.
21. Woods, S., K. Dautenhahn, and J. Schultz, *Child and adults' perspectives on robot appearance*, in *AISB05 Symposium on robot companion*2005, SSAISB: Hatfield, UK. p. 126-132.
22. Beck, A., et al. *Interpretation of Emotional Body Language Displayed by Robots*. in *Affine 2010*. Firenze, Italy: ACM.