

An Android Head for Social-Emotional Intervention for Children with Autism Spectrum Conditions

Andra Adams and Peter Robinson

University of Cambridge Computer Laboratory
{Andra.Adams,Peter.Robinson}@cl.cam.ac.uk

Abstract. Many children with autism spectrum conditions (ASC) have difficulties recognizing emotions from facial expressions. Behavioural interventions have attempted to address this issue but their drawbacks have prompted the exploration of new intervention strategies. Robots have proven to be an engaging and effective possibility. Our work will investigate the use of a facially-expressive android head as a social partner for children with ASC. The main goal of this research is to improve the emotion recognition capabilities of the children through observation, imitation and control of facial expressions on the android.

1 Introduction

Autism spectrum conditions (ASC) are characterized by difficulties in social interaction and communication, as well as repetitive behaviour and narrow interests [20]. The prevalence of diagnosed ASC has shown a steady increase over the past 40 years [6]. New technologies offer interesting possibilities for intervention techniques for children with ASC. In particular, this work investigates the use of a facially-expressive android head in an intervention to improve emotion recognition capabilities in children with ASC. Over a series of sessions, the children will both observe, imitate and control the facial expressions of the robot to enact various emotions.

2 Background

This interdisciplinary work builds on research in three main areas: autism spectrum conditions, human-robot interaction, and affective computing.

2.1 Autism Spectrum Conditions

Difficulties with facial expressions in ASC. Of particular interest to this work are the findings that children with ASC have difficulties with face processing tasks, as usefully summarized by Tanaka et al [28]. As a result, children with ASC do not become “face experts” in the same way that neurotypical children do, particularly since facial expressions are fleeting and unpredictable in

real-time social scenarios. We attempt to address this issue by providing children with ASC an opportunity to practise facial expression recognition in a more predictable setting than typical social environments.

The value of imitation as an intervention strategy. Even in individuals without ASC, imitation plays an important role in cognitive development [19]. Infants mimic their parents and others as a means of developing their social and communication skills. Not surprisingly, intervention techniques that employ reflection, imitation and synchronicity have been successful in improving the social behaviour of children with ASC [17,11].

Our research will use imitation in two ways. Firstly, the child is asked to imitate the facial expressions of a particular emotion that it has observed on the robot's face. This encourages the development of the child's own imitation capabilities. Secondly, as the child attempts to mimic what he/she has previously seen on the robot, the robot itself will be mimicking the facial expressions of the child. This serves both to show the child what his/her facial expressions look like and to encourage further social interaction.

Emotion recognition intervention strategies for children with ASC.

Emotion recognition plays a crucial role in effective social interaction. Our ability to attribute emotions to others from their outward appearance (facial, gestural, vocal) is often referred to as "mind-reading" [3], an ability which allows us to discern the mental states of others to better predict their beliefs, values, emotions and intentions. In particular, research has suggested that facial expressions are likely the most important source of affective information in humans [2].

Behavioural interventions targeting emotion recognition are common. These interventions typically consist of a few hours of training every week over a period of several months, and usually employ a teacher, therapist or parent that is familiar to the child [8]. But behavioural interventions are socially exhausting and often focus only on basic emotions since time constraints prevent further exploration [16]. New computer-based technologies can help to solve both of these issues without compromising the interactivity of a therapy session. Furthermore, computer-based technologies are often inherently motivating to children with ASC [4] and allow them to work at their own pace with as much or as little repetition as desired. Examples include the interactive emotional taxonomy of the Mind Reading DVD [5], the social scenarios of The Transporters series [12] and many others [27,15].

One of the important goals of any emotion recognition intervention is to ensure that the recognition of facial expressions and emotions is generalizable to unfamiliar scenarios. In this respect, there is an important difference between explicitly-taught and implicitly-taught interventions. Explicitly-taught interventions, like the Mind Reading DVD, are devoid of social context and confine the recognition of emotions to simplified scenarios. Implicitly-taught interventions, like The Transporters series, situate emotions in a relevant social context and are therefore more likely to generalize to real-world social scenarios.

2.2 Human-Robot Interaction

Robots for children with ASC. Human-robot interaction is a promising area for autism therapy. Robots are simplified, more predictable versions of human beings and research suggests that for many children with ASC, robots are more appealing and engaging than their human counterparts [26].

It is the fine balance between predictability and autonomy that makes robots promising social partners for children with ASC. Inanimate toys lack the ability to prompt children with ASC to break out of their repetitive behaviours and engage socially. On the opposite end of the spectrum, human peers behave far too unpredictably for a child with ASC, inducing stress and anxiety during social interaction. The controllable autonomy of robots offers the best of both worlds. Slowly increasing the autonomy of the robot challenges the child with ASC to engage in a variety of social interaction scenarios without pushing the child too far out of his/her comfort zone.

Robots vs software avatars. The fact that robots have a physical representation yet still have the controllable autonomy of software avatars make them ideal as learning tools for children with ASC. The physical representation of the robot provides engagement in physical space that cannot be achieved easily by a software avatar (e.g. physical proximity, touch, viewing angle) [9].

Androids vs simplistic robots. Simplistic robots, such as robot trucks and rolling spheres, have been successfully used in social interaction scenarios with children with ASC including basic motor imitation and “tag”-like games [18,10]. These simple robots often take the form of common toys to increase familiarity and mitigate adverse reactions from the children. The simple animism that is added to these formerly inanimate objects has been used to help with turn-taking, eye gaze, self-initiated behaviour, imitation and shared attention [10].

However, robots with simplistic forms do not have the features required to teach more complex social interactions. For example, teaching emotion recognition from facial expressions requires the robot to have basic facial features. These facial features can either be highly realistic or abstracted and cartoon-like. Research has suggested that children with ASC tend to be more comfortable with non-humanoid feature-less robots [23], which encourages the use of simple cartoon-like facial features. Yet the difficulty with generalization experienced by children with ASC suggests to us that more realistic human-like facial features will result in greater transferability of recognition rates to real human faces.

The FACE project at the University of Pisa is currently the only autism therapy project using a realistic android head [21]. Its aim is to improve the social and emotive capabilities of children with autism by using the android as an interface between the therapist and the child with ASC. To date, FACE is able to express and recognize Ekman’s six basic emotions at various levels of intensity, and is controlled by both the therapist and the child via a traditional



Fig. 1. Various facial expressions displayed on our android head from Hanson Robotics

screen-keyboard-mouse interface. Preliminary studies suggest that children with ASC are interested in interacting socially with the android, including spontaneous imitation of its head and facial movements. This successful social interaction with FACE suggests that realistic androids have an important role to play in autism intervention techniques. Our realistic android is shown in Figure 1.

2.3 Affective Computing

Selecting a corpus of emotionally-expressive data. There are many factors to consider when selecting a corpus of emotionally-expressive facial data. Naturalistic data is preferred to acted data since the timings and motions differ between the two and hence naturalistic data is more representative of the real-world [30]. However naturalistic data is extremely difficult to collect and label [1]. Corpora that include complex emotions are more desirable than those that contain only Ekman’s six basic emotions [24]. Furthermore, moving images are preferred to still images since humans are best able to recognize emotions through facial movement [7]. Lastly, facial appearance (age, sex, ethnicity) has been shown to shape the perception of emotional expressions [14], and therefore databases that feature multiple subjects are preferred to single-subject databases.

For our proposed work, the videos from the Mind Reading DVD have been selected as a reasonable corpus of emotional expression. The Mind Reading collection consists of more than 2400 videos depicting 412 different emotions with at least six different actors portraying each emotion. The 412 emotions are presented in a hierarchical taxonomy and have been validated by a panel of independent judges [13]. Although the emotional expressions in these videos are acted and not naturalistic, the wide variety of emotions and the use of multiple actors make this corpus an attractive choice.

Creating realistic expressions on the robot. It is essential that the robot accurately mimics the facial expressions in the Mind Reading videos. We use the FaceTracker software [25] to locate 66 feature points on the face which are then converted into motor movements for the robot, as shown in Figure 2.

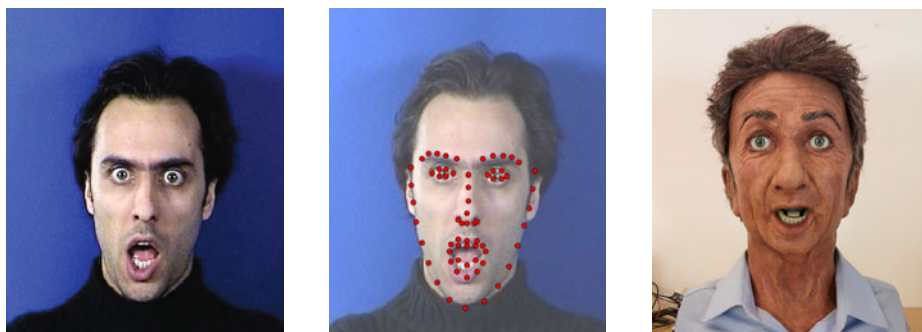


Fig. 2. Video images (left) are analyzed to track facial feature points (centre) which are converted to motor movements on the robot (right)

3 Methodology

We designed an intervention to assess the effect of an android head on the emotion recognition capabilities of children with ASC. Intended to span several months, the intervention begins with an initial familiarization period where the child is introduced to the robot in a comfortable, non-threatening manner. A therapist, teacher or parent is present to demonstrate the capabilities of the robot. Next, the child begins the two phases of the intervention: playback and imitation of the Mind Reading videos, and social interaction during a card game.

3.1 Playback and Imitation of the Mind Reading Videos

The first phase of the intervention is intended to expose the child to a wide variety of emotional facial expressions. An emotion is selected from the 412 emotions from the Mind Reading DVD and one of the six videos for that emotion is “acted out” by the robot while the child observes. The child is then asked to recreate the facial expressions of that emotion on the robot. To control the motion of the robot, the child’s own facial actions are used. A video camera pointed at the child’s face captures his/her facial expressions and this information is in turn fed back into the robot to control its facial movements.

This type of control is useful in three major ways. Firstly, it transforms the child from passive to active learner by engaging the child in an interactive exchange with the robot. Secondly, it encourages the child to look at all parts of the robot’s face rather than only the mouth region, as many children with ASC are prone to do [22]. Thirdly and most importantly, in order to control the robot, the child must first learn to imitate the facial expressions of the robot on his/her own face. As previously discussed, this use of imitation has great value in improving social interaction for children with ASC.

The child’s imitation is captured and evaluated based on accuracy (correctness of facial actions) and timeliness (ordering and timing between facial actions).

Feedback is given to the child both from the teacher/parent/therapist and from the algorithmic assessment of the captured video.

3.2 Social Interaction during a Card Game

Given the relative success of The Transporters series over the Mind Reading DVD [4], it is evident that situating facial expressions within a social context is an important factor for emotion recognition. Therefore in the second phase of the intervention, the child will play a card game such as “War” with the robot and the social context of the game will dictate the robot’s facial responses. Frustration, confusion, happiness, thinking, interest, sadness and many other emotions will be elicited by the card game and the robot will respond with appropriate facial expressions. The teacher/parent/therapist can interrupt the game at any point to encourage the child to consider the current mental state of the robot.

3.3 Evaluation of Intervention

The intervention will be evaluated through performance tasks completed by the children before and after the intervention with the robot. The children will be asked to label the emotions observed in videos from three sources:

- **Mind Reading DVD:** videos of acted emotions (some previously seen in the intervention and some unfamiliar), each portrayed by either a human actor or the robot
- **Recordings from the card game:** videos of the game, each showing either a child or the robot with surrounding social context
- **The Transporters series:** video clips of novel social situations with unfamiliar characters

As with the evaluation of The Transporters series [12], each video will be accompanied by three possible labels: the correct emotion, an emotion of opposite valence, and an emotion of the same valence. Our selection of performance tasks allows us to evaluate the children’s emotion recognition skills as well as their ability to generalize these skills to unfamiliar situations.

Furthermore, five groups of children will participate in the study to control for other factors and to compare against similar intervention strategies: an ASC robot intervention group, an ASC software avatar intervention group, an ASC Mind Reading intervention group, an ASC control group and a neurotypical control group. The ASC robot intervention group will participate in the intervention as described above, the ASC cartoon avatar intervention group will participate in the intervention described above but a software avatar will substitute for the robot, the ASC Mind Reading intervention group will use the Mind Reading DVD, and the two control groups will not participate in any intervention.

4 Conclusion

Our work will investigate the use of a facially-expressive android head in improving emotion recognition capabilities in children with ASC. We have designed an

intervention in which children are encouraged to observe, interact with and imitate the android, giving them valuable practice in attributing emotions to facial expressions. We have proposed detailed performance tasks for evaluating the success of this intervention technique, particularly concerning generalization of emotion recognition skills to real-world social situations.

The proposed work with the robot could be extended to other possible interventions, including addressing non-verbal social cues such as gaze-direction and turn-taking, or providing practice with social context via emotion elicitation activities.

Acknowledgements. This work is generously supported by the Gates Cambridge Trust. The sample video frame in Figure 2 uses the MMI-Facial Expression Database [29].

References

1. Afzal, S., Robinson, P.: Natural affect data - collection & annotation in a learning context. In: ACII 2009, pp. 1–7. IEEE, Los Alamitos (2009)
2. Ambady, N., Rosenthal, R.: Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin* 111(2) (1992)
3. Baron-Cohen, S.: *Mindblindness: an essay on autism and theory of mind*. MIT Press, Cambridge (2001)
4. Baron-Cohen, S., Golan, O., Ashwin, E.: Can emotion recognition be taught to children with autism spectrum conditions? *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535), 3567 (2009)
5. Baron-Cohen, S., Golan, O., Wheelwright, S., Hill, J.J.: *Mind reading: the interactive guide to emotions* (2004)
6. Baron-Cohen, S., Scott, F.J., Allison, C., Williams, J., Bolton, P., Matthews, F.E., Brayne, C.: Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry* 194(6), 500 (2009)
7. Bassili, J.N.: Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance* 4(3), 373 (1978)
8. Bauminger, N.: The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention outcomes. *Journal of Autism and Developmental Disorders* 32(4), 283–298 (2002)
9. Breazeal, C., Brooks, R.: Robot emotion: A functional perspective. *Who Needs Emotions*, 271–210 (2005)
10. Dautenhahn, K., Werry, I.: Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition* 12(1), 1–35 (2004)
11. Dawson, G., Adams, A.: Imitation and social responsiveness in autistic children. *Journal of Abnormal Child Psychology* 12(2), 209–226 (1984)
12. Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., Baron-Cohen, S.: Enhancing emotion recognition in children with autism spectrum conditions: an intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders* 40(3), 269–279 (2010)
13. Golan, O., Baron-Cohen, S.: Systemizing empathy: Teaching adults with asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology* 18(02), 591–617 (2006)

14. Hess, U., Adams, R.B., Kleck, R.E.: The face is not an empty canvas: how facial expressions interact with facial appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535), 3497 (2009)
15. Hopkins, I.M., Gower, M.W., Perez, T.A., Smith, D.S., Amthor, F.R., Casey Wimsatt, F., Biasini, F.J.: Avatar assistant: Improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders*, 1–13 (2011)
16. Howlin, P., Baron-Cohen, S., Hadwin, J.: Teaching children with autism to mind-read: A practical guide for teachers and parents. J. Wiley & Sons, Chichester (1999)
17. Ingersoll, B., Gergans, S.: The effect of a parent-implemented imitation intervention on spontaneous imitation skills in young children with autism. *Research in Developmental Disabilities* 28(2), 163–175 (2007)
18. Michaud, F., Th  berge-Turmel, C.: Mobile robotic toys and autism. *Socially Intelligent Agents*, pp. 125–132 (2002)
19. Nadel, J., Gu  rini, C., Pez  , A., Rivet, C.: The evolving nature of imitation as a format for communication (1999)
20. American Psychiatric Association Task Force on DSM-IV. In: *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. Amer. Psychiatric Pub. Inc., Washington (2000)
21. Pioggia, G., Sica, M.L., Ferro, M., Iglizzo, R., Muratori, F., Ahluwalia, A., De Rossi, D.: Human-robot interaction in autism: FACE, an android-based social therapy. In: *The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 605–612. IEEE, Los Alamitos (2007)
22. Riby, D.M., Doherty-Sneddon, G., Bruce, V.: The eyes or the mouth? Feature salience and unfamiliar face processing in williams syndrome and autism. *The Quarterly Journal of Experimental Psychology* 62(1), 189–203 (2009)
23. Robins, B., Dautenhahn, K., Dubowski, J.: Does appearance matter in the interaction of children with autism with a humanoid robot? *Interaction Studies* 7(3), 509–542 (2006)
24. Rozin, P., Cohen, A.B.: High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of americans. *Emotion* 3(1), 68 (2003)
25. Saragih, J.M., Lucey, S., Cohn, J.F.: Face alignment through subspace constrained mean-shifts. In: *2009 IEEE 12th International Conference on Computer Vision*, pp. 1034–1041. IEEE, Los Alamitos (2009)
26. Scassellati, B.: How social robots will help us to diagnose, treat, and understand autism. *Robotics Research*, 552–563 (2007)
27. Silver, M., Oakes, P.: Evaluation of a new computer intervention to teach people with autism or asperger syndrome to recognize and predict emotions in others. *Autism* 5(3), 299 (2001)
28. Tanaka, J.W., Wolf, J.M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., Brown, C., Stahl, S., Kaiser, M.D., Schultz, R.T.: Using computerized games to teach face recognition skills to children with autism spectrum disorder: the Let’s Face It! program. *Journal of Child Psychology and Psychiatry* 51(8), 944–952 (2010)
29. Valstar, M.F., Pantic, M.: Induced disgust, happiness and surprise: an addition to the MMI facial expression database. In: *International Language Resources and Evaluation Conference* (May 2010)
30. Valstar, M.F., Pantic, M., Ambadar, Z., Cohn, J.F.: Spontaneous vs. posed facial behavior: automatic analysis of brow actions. In: *Proceedings of the 8th International Conference on Multimodal Interfaces, ICMI 2006*, pp. 162–170. ACM, New York (2006)