Building a game scenario to encourage children with autism to recognize and label emotions using a humanoid robot

5 authors, including:

Sandra Costa
University of Minho
26 PUBLICATIONS   **572** CITATIONS

SEE PROFILE

Filomena O. Soares
University of Minho
323 PUBLICATIONS   **2,361** CITATIONS

SEE PROFILE

Ana Paula Silva Pereira
University of Minho
98 PUBLICATIONS   **366** CITATIONS

SEE PROFILE

Cristina Peixoto Santos
University of Minho
288 PUBLICATIONS   **2,844** CITATIONS

SEE PROFILE

Building a Game Scenario to Encourage Children with Autism to Recognize and Label Emotions using a Humanoid Robot

Sandra Costa[1], Filomena Soares[1], Ana Paula Pereira[2], Cristina Santos[1] and Antoine Hiolle[3]

Abstract— This paper presents an exploratory study in which children with autism interact with ZECA (Zeno Engaging Children with Autism). ZECA is a humanoid robot with a face covered with a material allowing the display of varied facial expressions. The study investigates a novel scenario for robot-assisted play, to help promoting labelling of emotions by children with autism spectrum disorders (ASD). The study was performed during three sessions with two boys diagnosed with ASD. The results obtained from the analysis of the children's behaviours while interacting with ZECA helped us improve several aspects of our game scenario such as the technical specificities of the game and its dynamics, and the experimental setup. The software produced for this study allows the robot to autonomously identify the answers of the child during the session. This automatic identification helped the fluidity of the game and freed the experimenter to participate in triadic interactions with the child. The evaluation of the game scenario that will be used in a future study was the main goal of this pilot study, rather than to quantify and evaluate the performance of the children. Overall, this exploratory study in teaching children about labelling emotions using a humanoid robot embedded in a game scenario demonstrated the possible positive outcomes this child-robot interaction can produce and highlighted the issues regarding data collection and their analysis that will inform future studies.

I. INTRODUCTION

According to the current criteria in the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition), Autism Spectrum Disorders (ASD) are characterized by repetitive patterns of behaviour, restricted activities or interests and impairments in social communication [1]. These impairments in social communication in children are mostly observed in their difficulty to respond to social stimuli, to imitate behaviours, to recognize and understand mental states in themselves and in others [2], [3]. These differences clearly influence the adaptation of children with ASD to their natural contexts with implications to their cognitive, linguistic, and emotional skills [4].

The Robotica-Autismo project aims to use social robots as promoters of social emotional development in children with ASD. As a central tool to our research, we employ a humanoid robot with the ability to display facial expressions, to promote social interaction, communication and emotion recognition. This tool fits into the category of emotional robotics [5], defined as robotic systems that rely on the detection, synthesis and production of affective behaviours to allow visual recognition of the presented emotion.

The research question motivating our research can be stated as follows: Can robots help developing emotion recognition skills in children with ASD? To address such issues, we intend to develop a research design that can adequately capture the dynamics inherent to the learning process of these children. However, before starting a long study with a target group composed by a large sample, some constraints need to be assessed. This procedure needs to be done having in mind the target group of this study which can be very unpredictable. This paper presents the results from a pilot study with two participants, the main goal of which was to test our game scenario with the robot in which the children had to label facial expressions and matching gestures. This labelling was achieved by association images of emotional faces to the emotions the robot had displayed.

In Section II we present other research projects which also use human-robot interactions to help children with ASD to recognize and label emotions. Section III features the procedures used during the experiments. Section IV and V contain the results and the discussion, respectively. Conclusions and future work are presented in Section VI.

II. BACKGROUND

Some projects dedicated their research to the specific topic of the design of emotional expression for robots used in human-robot interaction (HRI). However, only a limited number of projects focuses specifically on the use of robots with children with ASD as promoters of the recognition of facial expressions and emotions.

WikiTherapist [6] is a platform intended to be used by therapists to develop intervention programs for individuals with ASD. It aimed to promote incorporated interaction, through interfaces able to infer intention. This intention emerged from human being's natural movements. The authors described a framework which enabled the expression and interpretation of patterns of emotional movement. This project contributes to the Robotica-Autismo approach especially regarding the modelling and the analysis tools to design scenarios with robots. However, the robot used in this project was a NAO robot from Aldebaran Robotics which was able to perform gestures with its hands and head. It expresses emotions by performing head movements and changing the colour of the eyes. These movements may not be enough for children with ASD to acquire emotion recog-

[1]Sandra Costa, Filomena Soares and Cristina Santos are with R&D Centre Algoritmi, School of Engineering, University of Minho, Portugal `scosta@dei.uminho.pt`, `fsoares@dei.uminho.pt`, `cristina@dei.uminho.pt`

[2]Ana Paula Pereira is with the Institute on Children Studies, University of Minho, Portugal. `appereira@ie.uminho.pt`

[3]Antoine Hiolle is with the Embodied Emotion, Cognition and (Inter-)Action Lab, University of Hertfordshire, United Kingdom. `a.1.hiolle@herts.ac.uk`

nition skills. Children with ASD need the facial expressions to be strong and marked so they can perceive them as such [7]. The authors focused their attention on the therapists' point of view as having the main role in the intervention.

The humanoid robot FACE [8] was built to allow children with ASD to deal with expressive and emotional information. This adaptive platform integrates information derived from sensors used on children and in the environment. The expressions and movements of FACE were modelled to be harmonized with the feelings of the user. Together with a shirt with sensors, video cameras and an eye tracking hat, the researchers expected to assess whether children with ASD could learn empathy. However, this project has an important limitation. Some children with ASD have problems with tactile interactions, especially when they are the object thereof. Hence, the possibility of having wearable sensors on the children may not be an option. The sensors around the room can be really useful though, as is an infrared camera that measures the temperature of the child while interacting with the robot. In contrast with this work, our approach will use video analysis of the children's behaviours to evaluate their progress and interaction between the robot and the experimenter.

The Bandit robot [9] intended to facilitate human-robot interactions in a natural way, improving child's interactions skills. The authors of this project found that different embodiments of robots could interact safely with children with ASD and that verbal children showed more interest. Our team considers that it is important to correctly manage the child's expectations, and to enhance the complexity of the robot's social behaviours. This applies to whether the children are low- or high-functioning and verbal or non-verbal. However, considering that the Bandit robot only shows movements with its mouth and eyebrows, it would be certainly difficult to produce more than the corresponding expressions to happiness and sadness. Preferably, the robot should be able to display the basic emotions defined by Ekman, since they are held to be so because they are universally recognized and expressed in the same way [10]. In addition, the authors focused on the robot control architecture having in mind the adaptation for use by non-roboticists engaged in ASD therapy, while we focus on the children's performance and its interaction with the robot.

Probo [11] is an animal-like robot, designed to act as a social interface. The authors used Probo as a platform to study HRI and it was capable of performing basic facial expressions. The authors described Probo's ability to express emotional states. In their opinion, a better recognition of the robot's facial expressions contributes to the general social acceptance. This opinion is shared with the authors of this paper. In addition, the recognition of the facial expressions is important for an effective non-verbal communication between a human and a robot. This project emphasizes facial features, which play an important role regarding the expression of emotions. Additionally to these main features, our project takes into consideration the addition of gestures as an important input for a faster recognition, as a result from

[12].

In contrast with the projects presented above, we consider that a study involving a robot with the characteristics of ZECA may bring interesting insights about how the skill of recognizing and labelling emotions may be promoted in children with ASD. Additionally, we consider the addition of gestures compliant with the represented emotions will help the children to use them in context. Based on the literature and on the opinions of several professionals, the motivation of the game and corresponding setup presented in this paper focus on the design and testing of the game scenario to be used in a future study with a larger sample.

III. PROCEDURE

The robot we used differs greatly from robots used in other designs due to the face being covered with a polymeric material called Frubber, giving it the ability to display varied facial expressions (Fig. 1). Besides this expressive face, this humanoid robot, developed by Hanson Robotics, possesses a walking body (with 31 degrees of freedom in total), a loud speaker on the chest, and several sensors including two HD cameras in each eye [13].

Fig. 1.   ZECA, humanoid robot produced by Hanson Robotics.

The software developed allowed the robot to autonomously identify the answers of the child during the session. This automation was considered necessary to help the fluidity of the game and to free the experimenter to interact with the child.

Two different tasks were tested in this pilot study. The first task, called Performance task, was chosen with the help of special education needs teachers and will be used in the future study to evaluate the skill level of children in labelling emotions. The comparison of the results of this task are not going to be presented for two reasons. On one hand because the programmed number of sessions was not enough to acquire a new skill, based on the professionals' expertise, and on the other hand because the goal of this pilot study was not to evaluate the performance of the children, but the task itself and its potential benefits and shortcomings.

The second task, from now on called Recognize, was presented to the children individually. Each task will be presented in the following subsections.

A. Performance Task

This task has the final goal of evaluating the emotion labelling's skill of the children. This task consisted in matching cards on which a man or a woman is showing one of five different emotions (happiness, sadness, anger, surprise, and fear). These cards were matched with cards with PECS (Picture Exchange Communication System) representing the same emotions. Examples of the cards showed to the children are presented in Fig. 2.
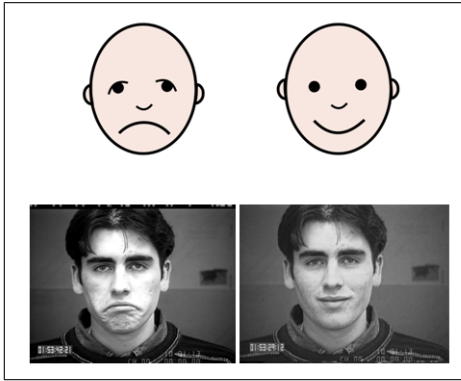

Fig. 2. Example of images used in the performance task. The top of the figure shows the PECS cards which were matched to the images of faces of a man shown in the bottom of the figure. These cards represent the emotions sadness and joy.

The choice of the four persons displayed in the pictures (2 men and 2 women) was based on the work of [14], [15]. This database was released for the purpose of promoting research into automatically detecting individual facial expressions. The five PECS cards were presented to the child at the same time on a board, attached to it using Velcro. Five empty spaces under the PECS cards were available, and the experimenter delivered one card with the picture of the man or the woman, and prompted the child to match the card he had in his hand with the ones on the board. Once the child managed to correctly match the cards, the experimenter gave him another one, until they were over.

B. Recognize

The task Recognize consisted in the robot first displaying a facial expression and its associated gestures (as a body posture), representing one of the five basic emotions: fear, joy, sadness, surprise, and anger. The child is then prompted to identify the emotion associated with the facial expression. The facial expressions and gestures of the robot were previously evaluated by a group of 61 adults and a group of 42 children normally developed with quite good efficiency (between 70 and 93% of correct labelling) [12]. The child answers by selecting one of five rackets presented in front of him and showing it to the robot. Each racket featured a picture with a face representing an emotion and its corresponding label. Additionally, each racket had a

Quick Response (QR) code which was used to automatically identify the emotion (Fig. 3). This QR code was then read by one of the HD cameras of the robot. The option of having the experimenter assessing the child's answer was considered. However, the authors considered that it was important for this autonomous process to exist to allow the experimenter to interact freely with the child.

The experiment started with the robot prompting the child to match its facial expression and gesture to an emotion represented in 5 different rackets. The child answered, choosing the corresponding racket. When the child answered correctly, the robot gave him a reward based on the type of reinforcement his/her teacher identified as his/her favourite (either movement, verbal, or both). If the answer was incorrect, the robot shook its head and said, for example "Ups. Pay attention. Let's try another one!".

Fig. 3. Rackets used in the study for the children to answer the prompts of the robot. Each racket features a picture of a face, a label, and QR code corresponding to the emotion.

The images displayed on the rackets were chosen considering the opinion from professionals working in special education. Four options were presented: PECS, experimenter, ZECA, and unknown persons. The first option was discarded because even though these cards are normally used with these children to develop other types of skills, they present the difficulty of generalizing the labelling of emotions to human beings. Using an image of the experimenter could be an advantage , and the generalization could be easier, but the fact that the experimenter was also in the room could lead the child to compare the racket to the experimenter and not to the robot. We excluded this option to try to prevent all sources of distractions. We discarded the third option as well not to hinder potential generalisation. Thus, the chosen option was the images with unknown persons, so it could be easier for the children to generalize to another human being. These images were also tested and validated using an on-line questionnaire with a group of 76 adults (with the age range: M = 28.26; SD = 8.31). The ratings given by the participants showed a good accuracy level (between 96 and 100% of correct labelling). Specifically, the recognition rates were: fear - 100%, joy - 100%, sadness - 100%, surprise - 96.05%, and anger - 98.68%.

C. Participants

The participants in this pilot study were boys with ASD aged fourteen and sixteen years old. The participants were high-functioning, according to their diagnosis. The experimenter was in the room to introduce the robot, and to intervene in case of difficulties. She was also involved in the activity as a facilitator of the interaction, providing guidance and ensuring that the children did not become agitated or damage the robot during the activity. A signed informed consent form was obtained from the parents of each child.

D. Experimental setting

During this study, the sessions took place in an individual context, encouraging triadic relationships between children, the experimenter and the robot (Fig. 4). The produced videos were analysed using the specialized software The Observer XT from Noldus [16] to quantify predetermined behaviours by the children during the sessions. The experiments were carried out by the first author.
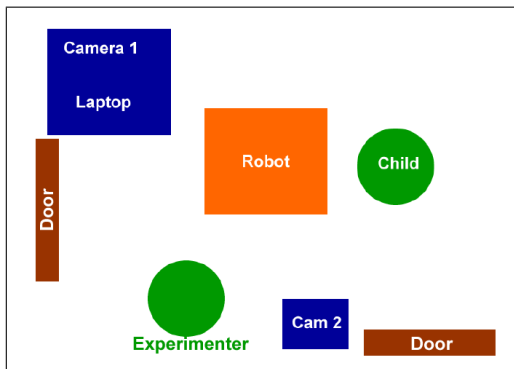

Fig. 4. Room setup which comprises two cameras to record two different angles of the interaction, and one laptop. The robot in the centre of the room forms a triangle with the child and the experimenter, promoting a triadic interaction.

E. Evaluation Tools

The coded behaviours were the following: tactile interactions, prompts, and answers. For tactile interaction, the coders marked whether the child showed the behaviour spontaneously or whether the behaviour was prompted by the experimenter. If the child was, for example, touching the robot for no specific reason, the behaviour was classified as spontaneous. Tactile interaction was also classified as gentle or harsh, according to the pressure applied from the child on the robot. When the child exhibited behaviours that were not specified in our list, they were not coded. The selection of the target behaviours was done according to its relevance (regarding engagement, and interactive behaviour), and the feasibility in identifying them in the recorded data. Additionally, as a quantitative measure, we were interested in recording the children's response time (from the end of the prompt until the child showed the racket).

Besides the quantitative evaluation previously described, a qualitative analysis was done. This analysis had two sources: observation of the videos, and the room layout. From the first one, we wished to identify particular events such as the first reaction of the child to the robot, and if the child got out of the room before the end of the session. Concerning the last topic, since the HD camera of the robot is used to get the answers from the child, we wanted to assess whether the room lighting and the configuration used were suitable for this automated answer recognition and if the height of the robot was appropriate for the interaction with the child.

IV. RESULTS

Fig. 5a shows that Participant 1 gave more incorrect answers than correct ones, whereas Participant 2 gave more incorrect answers in Session 2 only, as can be seen in Fig. 5b. During Session 3, Participant 2 correctly answered more than half of the prompts, and during the first session the number of correct and incorrect answers were equal. The caption of Fig. 5 presents the session's durations varying between approximately 250 to 1000 seconds. This variation was caused by the duration of the Performance Task.

Fig. 6 shows how tactile interactions with the robot evolved during the sessions. Only gentle touches were observed by the children and the prompts from the experimenter were kept to a minimum. Participant 1 showed a lot of interest in touching the robot, exploring it during the game (Fig. 6 a). The robot's body parts touched more often were the face, hands, feet and chest. Participant 2 was more involved in the game and he touched the robot more often in the second session (Fig. 6 b).

Tables I and II present the children's mean (M) response time, and the standard deviation (SD) of unsuccessful and successful answers given in each session. These values represent the duration between the time the robot gave the prompt until the child showed the racket. Both participants took more time answering to the prompt in Session 3. Participant 1 was usually faster to answer the prompt from the robot than Participant 2. If we look individually to each participant's data, we verify that Participant 1, only improved his performance regarding the display of anger and happiness. Participant 2 improved his performance in labelling fear, happiness, sadness, and surprise.

TABLE I

CHILDREN'S MEAN RESPONSE TIME IN SECONDS FOR UNSUCCESSFUL ANSWERS. (SD)

	Session 1	Session 2	Session 3
Participant 1	6.54 (7.72)	9.22 (6.95)	15.88 (10.03)
Participant 2	9.06 (10.37)	30.93 (22.49)	40.31 (27.8)

When we observe the number of attempts for the child to show the racket, we verified that even with the correct lighting in the room, sometimes the QR code was not read because the child put the racket too close to the robot preventing the camera to get the entire QR code and then to make the reading. This caused the experimenter to interfere in the session to help the child to show the racket.
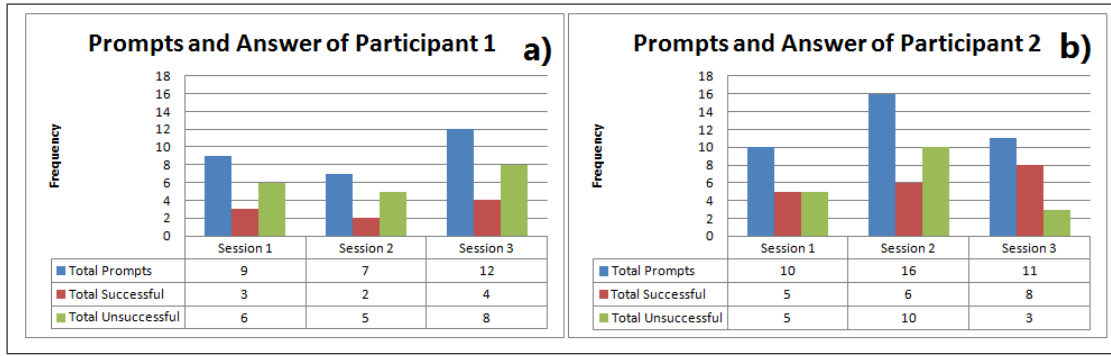
Fig. 5. Number of Prompts from the robot, Successful Answers, and Unsuccessful Answers for a) Participant 1 - Total Session Time (sec): S1: 243.84; S2: 334.96; S3: 593.04; b) Participant 2 - Total Session Time (sec): S1: 258.72; S2: 996.96; S3: 1010.96.
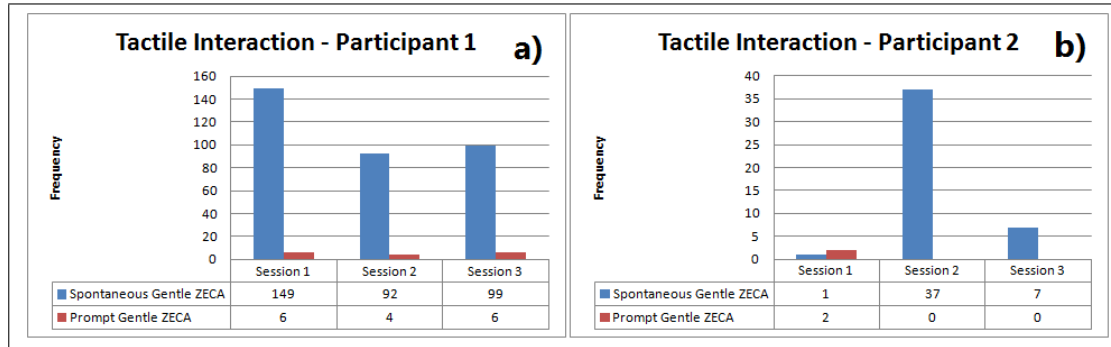

Fig. 6. Number of times the participants touched the robot, either spontaneously or prompted by the experimenter.

TABLE II
CHILDREN'S MEAN RESPONSE TIME IN SECONDS FOR SUCCESSFUL ANSWERS. (SD)

	Session 1	Session 2	Session 3
Participant 1	10.65 (9.83)	4.88 (1.81)	25.03 (15.6)
Participant 2	6.58 (6.95)	39.97 (17.09)	50.84 (20.32)

The first qualitative remark we would like to present is the first reaction of the children to the robot in the first session. Both children were specifically interested in the face of the robot, touching it repeatedly and always in a gentle way. Participant 1 also touched the robot on the chest several times. None of the children abandoned the room, or got up of his chair during the sessions indicating that, even while interacting with a new tool they were interested in the new object. Since the robot's input to receive the answer from the child was based on image processing, one particularly important aspect is the room lighting. Only one modification to the room had to be performed so the QR Code could be identified by the cameras, and this involved covering one of the windows on the door with a blanket. This modification did not influence the sessions since it was made before the first session, and the conditions were the same until the end. The first session was also useful to evaluate if the height of the robot was adequate for the interaction with the child. We noticed that the child had to look up often, so in the next sessions the robot was put lower so its head was at eye level

with the child, thus facilitating the interaction.

V. DISCUSSION

Even though the main goal of this study did not focus on the children's performance, it is interesting to notice that there is a knowledge gap appearing when we compare the results of both participants, giving us a margin to progress. The number of sessions, as suggested before, was not enough to improve children's accuracy on average. But individually, we have noticed some improvement in labelling specific emotions. The noticeable increase in the response time over the sessions might be related to the children thinking and considering all options they have available to them. Further investigation is going to be made regarding the relation between the correct answers after incorrect ones, their response time and whether the children manage to answer correctly at the beginning or at the end of the session.

The participants only touched ZECA gently, which might indicate that the children assumed that ZECA would react to a harsh touch, such as their friends would. However, more sessions are necessary to verify if the children would try different tactile interactions or even transfer it to the experimenter, as it was seen in related experiments [17].

With the observations from this pilot study, we were able to identify several aspects to improve our game scenario. These improvements will be implemented in the larger study we planned and they will help focusing on the children's performance.

The first version of the pre-test's task had four sets with

people displaying five different facial expressions. Presenting the four sets at the same time was too confusing for the children. Even presenting four sets separately was too tiring. To simplify the task, but still assessing the ability of the child to match facial expressions and emotions, the number of sets was reduced to two, one with a man and the other with a woman. One of the sets is going to be used as the pre-test and the other as a post-test, preventing the children from using the first association between the image and the emotion.

In the first session, the robot's support (80 cm high) was put on a table 75 cm high, causing the children to always have to look up towards the robot while seating on a chair, and then to look down to choose a racket. In the following sessions, the robot's support was put on a box at a height of 30 cm. This change places the head of the robot at eye level with the child during the session, facilitating the process of observing the facial expression, and choosing a racket.

In some cases, the QR code's reading failed due to the children's lack of gross motor coordination. To cater for failures in reading the QR code, the software was modified to accept inputs from a wireless numeric keypad used by the experimenter, matching numbers to the choice of the child to answer the prompt. This option is going only to be used for this purpose, since the authors believe it is important that the experimenter remains as free as possible to interact with the child when needed.

VI. CONCLUSIONS AND FUTURE WORK

Professionals keep struggling everyday because normally children with ASD need extra time for acquiring their skills and qualified professionals are not enough to attend these necessities. Our research project aims to use robotics social care for children with ASD. ASD are a group of lifelong disabilities affecting people in communicating and understanding social cues. Therefore, the use of robots to provide care for children from this spectrum through social interactions is eagerly awaited by professionals. One important remark is that this research does not intend to replace the work performed by professionals with children with ASD, but to provide a complementary tool to foster social interactions and emotion recognition using a humanoid robot. Being a pilot, this study had the primary goal to evaluate a game scenario to be used in the next study, and several improvements were needed and then implemented. This first contact of this robot and children with ASD allowed us to prepare the activities for children to develop their skills.

The future work includes the test of different game scenarios with the same robot. This larger study will include a sample of 45 children divided in three groups: experimental group with 15 children who will interact with the robot, control group 1 with 15 children who will perform the same tasks without the robot, and control group 2 with 15 children who only are going to perform the pre- and post-tests. We expect with these experiments to understand if and how the robot can be used as a valuable tool to develop skills of emotional labelling by children with ASD.

REFERENCES

[1] American Psychiatric Association: Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. edn. Autor, Washington, DC (2013)

[2] Clark, T.F., Winkielman, P., McIntosh, D.N.: Autism and the extraction of emotion from briefly presented facial expressions: Stumbling at the first step of empathy. Emotion **8**(6) (2008) 803

[3] Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., Szatmari, P.: Behavioral manifestations of autism in the first year of life. International Journal of Developmental Neuroscience **23**(2) (2005) 143–152

[4] Charman, T., Stone, W.L.: Social and communication development in autism spectrum disorders: Early identification, diagnosis, and intervention. Guilford Press (2006)

[5] Feil-Seifer, D., Mataric, M.J.: Defining socially assistive robotics. In: Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on, IEEE (2005) 465–468

[6] Gillesen, J., Boere, S., Barakova, E.: Wikitherapist. In: Proceedings of the 28th Annual European Conference on Cognitive Ergonomics, ACM (2010) 373–374

[7] Baron-Cohen, S.: The development of a theory of mind in autism: deviance and delay? Psychiatric Clinics of North America (1991)

[8] Mazzei, D., Billeci, L., Armato, A., Lazzeri, N., Cisternino, A., Pioggia, G., Igliozzi, R., Muratori, F., Ahluwalia, A., De Rossi, D.: The face of autism. In: RO-MAN, 2010 IEEE, IEEE (2010) 791–796

[9] Feil-Seifer, D., Mataric, M.J.: B 3 ia: A control architecture for autonomous robot-assisted behavior intervention for children with autism spectrum disorders. In: Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on, IEEE (2008) 328–333

[10] Ekman, P., Rosenberg, E.L.: What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press (1997)

[11] Saldien, J., Goris, K., Vanderborght, B., Vanderfaeillie, J., Lefeber, D.: Expressing emotions with the social robot probo. International Journal of Social Robotics **2**(4) (2010) 377–389

[12] Costa, S., Soares, F., Santos, C.: Facial expressions and gestures to convey emotions with a humanoid robot. In Herrmann, G., Pearson, M., Lenz, A., Bremner, P., Spiers, A., Leonards, U., eds.: Social Robotics. Volume 8239 of Lecture Notes in Computer Science. Springer International Publishing (2013) 542–551

[13] Hanson, D., Baurmann, S., Riccio, T., Margolin, R., Dockins, T., Tavares, M., Carpenter, K.: Zeno: a cognitive character. In: AI Magazine, and special Proc. of AAAI National Conference, Chicago. (2009)

[14] Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. In: Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, IEEE (2000) 46–53

[15] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, IEEE (2010) 94–101

[16] Noldus, L.: The observer: A software system for collection and analysis of observational data. Behavior Research Methods, Instruments, & Computers **23**(3) (1991) 415–429

[17] Costa, S., Lehmann, H., Robins, B., Dautenhahn, K., Soares, F.: where is your nose?-developing body awareness skills among children with autism using a humanoid robot. In: ACHI 2013, The Sixth International Conference on Advances in Computer-Human Interactions. (2013) 117–122