Proceedings of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment

Designing an Intelligent Virtual Agent for Social Communication in Autism

Sara Bernardini
Department of Informatics
King's College London
London, UK, WC2R 2LS
sara.bernardini@kcl.ac.uk

Kaśka Porayska-Pomsta
London Knowledge Lab
Institute of Education
London, UK, WC1N 3QS
k.porayska-pomsta@ioe.ac.uk

Harini Sampath
Center for IT in Education
IIIT-H
Hyderabad, India, 500 032
harini@research.iiit.ac.in

Abstract

This paper describes the Intelligent Engine (IE) of ECHOES, a serious game built for helping young children with Autism Spectrum Conditions acquire social communication skills. ECHOES IE's main component is an autonomous virtual agent that acts as a credible social partner for children with autism by engaging them in interactive learning activities. The other IE components are a user model, a drama manager and a social communication engine. We discuss how AI technology allows us to satisfy the requirements for the design of the agent and the learning activities that we identified through consultations with children and carers and a review of best practices for autism intervention. We present experimental results pertaining to the agent's effectiveness, which show encouraging improvements for a number of children.

1 Introduction

We present the *Intelligent Engine* (IE) of ECHOES, a serious game built for helping young children with Autism Spectrum Conditions (ASCs) acquire social communication skills (Porayska-Pomsta et al. 2012; 2013). Autism is a neuro-developmental disorder affecting three areas ("triad of impairments" (American Psychiatric Association 2000)): (i) *communication*: verbal and non-verbal language; (ii) *social interaction*: recognition and understanding other people's emotions and expression of own emotions; and (iii) *restricted or repetitive behaviours*: adaptation to novel environments. Our focus on *social communication* in this context is motivated by recent studies showing that this domain is regarded as the most challenging by ASCs children (Prizant et al. 2003), and support in this area is considered as the most desirable feature of technology-enhanced intervention (Putnam and Chong 2008). Social communication requires coordination and sharing of attention, intentions and emotions with others and a capacity for verbal and non-verbal reciprocal interaction.

The main component of the ECHOES IE is an *autonomous planning-based agent* that drives the decision making of the ECHOES virtual character, called Andy (Bernardini et al. 2012; Bernardini and Porayska-Pomsta 2013). Andy inhabits a virtual environment created for real-world use in schools and at home. A 42" multitouch LCD display with eye-gaze tracking facilitates the interaction between the child and the agent. The other components of the ECHOES IE are: (i) a *user model*; (ii) a *drama manager*; and (iii) a *social communication engine*. They complement the agent's architecture allowing Andy to interpret children's behaviours and to act as a credible social partner for them.

This paper is organised as follows. Sections 2 and 3 present related work and the pedagogical underpinnings of ECHOES. Sections 4 and 5 describe the design and the implementation of the ECHOES IE. Section 6 discusses the experimental results and Section 7 offers conclusions.

2 Assistive Technologies for Autism

Several studies show an affinity of ASC people with technology and computer-based training (Putnam and Chong 2008). Software programs offer the predictable and structured environment that is in line with this population's preference for routine and repetitive behaviours (Murray 1997). Virtual agents provide particular benefits to ASCs children (Parsons and Cobb 2011), who often find real social interactions stressful and unpredictable. Traditional educational settings, for example, are a challenge for autistic children. Social anxiety can be mitigated by artificial tutors that are capable of acting tirelessly, consistently and positively regardless of the child's behaviours. Artificial tutors can support individualised learning by meeting the individual children's needs: studies show that autistic children who were taught by a virtual human retain more information than they do in classrooms (Grynszpan, Martin, and Nadel 2008). Virtual reality may lead to generalisation because both role-play and practice of behaviours can facilitate transfer of the learned skills from the virtual to the real world (Parsons and Cobb 2011). Examples of agents for ASCs children include the life-sized virtual peer by Tartaro and Cassell (2008), which engages the children in collaborative narrative creation, the virtual tutor by Milne et al. (2010), which trains the children in responding to facial expressions, and the conversational agents used to teach *literacy skills* by Bosseler and Massaro (2003).

Despite the growing interest in the potential of artificial agents for autism intervention, the efforts have focused primarily on agents with little or no *autonomy*. Typically, virtual agents are either authored *a priori* or controlled by a practitioner through a control panel (e.g. Tartaro and Cassell (2008)). The Thinking Head (Milne et al. 2010), a 3-

D computer-animated talking head that teaches social skills by realistically portraying facial expressions, and Baldi and Timo (Bosseler and Massaro 2003), also talking heads for training language and speech, are the only projects that devote attention to autonomy in virtual reality. We argue that autonomous agents carry significant potential for autism intervention, because they can contribute to the intensive one-on-one support that ASCs children need while easing the demand for such support from practitioners and parents. Autonomous agents can complement traditional intervention by undertaking repetitive tasks and providing on-demand intervention and therefore leaving only the most complex aspects of face-to-face interventions to human practitioners.

3 Pedagogical Underpinnings of ECHOES

Our goal was to create an artificial social partner that could act credibly both as a peer and as a tutor for ASCs children and, as a result, deliver the educational and interpersonal support needed by these children to develop social communication skills. In designing this agent, we drew on best practices for autism intervention and consulted with users through two knowledge elicitation workshops involving thirty experienced practitioners and three high functioning teenagers with ASCs. Through storyboarding activities, group discussions and individual interviews, we assembled the requirements for the agent, which we further validated against SCERTS (Prizant et al. 2003), a comprehensive approach to social communication assessment and intervention in autism. SCERTS identifies the essential skills for successful social communication, which are encapsulated in three domains: (i) *Social Communication (SC)*: spontaneous and functional communication, emotional expression, and secure and trusting relationships with children and adults; (ii) *Emotional Regulation (ER)*: the ability to maintain a well-regulated emotional state to cope with everyday stress and be available for learning and interacting; and (iii) *Transactional Support (TS)*: the development and implementation of supports to help caregivers respond to the child's needs and interests, modify and adapt the environment, and provide tools to enhance learning. SCERTS breaks down each domain into a number of components and for each provides a detailed description of the objectives to be achieved, the strategies for intervention and the assessment criteria. We built on the this *operationalisation of social communication* in designing our agent's behaviours.

Learning Activities The interaction between the child and the agent is structured around twelve *learning activities*, which focus on social communication and, in particular on: (i) *Joint Attention:* child's ability to coordinate and share attention by looking toward people or shifting gaze between people and objects, share emotions by using facial expressions, express intentions, engage in turn-taking and participate in reciprocal social interactions by initiating/responding to bids for interaction; and (ii) *Symbol Use:* child's understanding of meaning expressed through conventional gestures, words, and sentences and child's ability to use nonverbal means and vocalisations to share intentions. The learning activities directly correspond to the intervention goals specified in the SCERTS framework, although some of them were adapted to fit the human-computer interaction context.

One of SCERTS principles is that in order to support shared attention learning activities need to share "an obvious unifying theme" (Prizant et al. 2006). Hence, all ECHOES activities take place in a "magic" garden populated by Andy and by interactive objects that behave in unusual ways, sometimes transforming into other objects when the agent or the child act upon them through specific touch gestures. For example, tapping the petals of a flower makes the flower become a floating bubble or a bouncy ball. We also follow the SCERTS philosophy advocating that learning activities need to be "meaningful and purposeful" (Prizant et al. 2006), in contrast with approaches in which the activities are task-based and skills are trained in a repetitive fashion and in isolation from a meaningful context. We designed two sets of activities: (i) *Goal-oriented* activities, with clear sequence of steps and an easily identifiable end-goal; and (ii) *Cooperative turn-taking* activities, with no clear end-goal and whose main objectives are social reciprocity, turn taking, and mutual enjoyment. Sorting a set of balls according to their colours and collecting all the flowers on the ground into a basket are examples of goal-oriented activities, while taking turns with the agent growing flowers by shaking a cloud that produces rain and throwing balls thorough a cloud so that they change colour constitute turn-taking activities. All activities are to be performed by Andy and the child in cooperation, with Andy assuming a more or less prominent role according to a particular activity's learning objective and the needs of the individual child. For example, if the goal is *learning-by-imitation*, Andy will adopt a leading role and will demonstrate different behaviours to the child. If the goal is *engaging-in-reciprocal-interaction*, Andy will wait to give the child an opportunity to initiate a bid for interaction, before initiating the interaction.

4 The ECHOES Autonomous Agent

Agent's Requirements: Based on SCERTS (Prizant et al. 2003, p. 309) and input from practitioners, we assembled our requirements for Andy:

1. **Role:** As a *tutor*, our agent needs to deliver *visual and organisational support* for: (a) "expanding and enhancing the development of a child's expressive communication system"; (b) "supporting a child's understanding of language as well as others' nonverbal behaviour"; and (c) "supporting a child's sense of organisation, activity structure, and sense of time". When acting as a *peer*, the agent needs to provide children with *interpersonal support* by: (a) accommodating the children's preference for structure and predictability, while fostering initiation, spontaneity, and self-determination; and (b) exposing the children to positive interactions with peers so that they can "benefit optimally from good language, social, and play models".
2. **Pedagogical Focus:** The agent's actions need to be manifestations of joint attention and symbolic use.
3. **Responsiveness:** Responsiveness should range from simple physical reactions to the ability to respond to the

child's changing needs and mental states. The agent should attune its emotional tone to that of the child to maintain emotional engagement with them and to provide them with positive feedback to promote the child's sense of self-efficacy and achievement.

4. **Style of Interaction:** The agent should exhibit an optimal interaction style, i.e. "one that provides enough structure to support a child's attentional focus, situational understanding, emotional regulation, and positive emotional experience, but that also fosters initiation, spontaneity, flexibility, problem-solving, and self-determination".

Agent's Architecture: Among the various domain-independent agent architectures proposed for building agents, FAtiMA (Dias and Paiva 2005; Aylett, Dias, and Paiva 2006) is well suited to fulfil our design requirements, because it integrates an affective appraisal system with a planning mechanism. A FAtiMA agent displays the reactive capabilities needed to obtain a responsive character (Requirement no. 3), the cognitive capabilities needed to provide the child with structured and goal-oriented activities (Requirement no. 1 and 4) and the socio-emotional competence necessary to help the child acquire social skills (Requirement no. 4). The emotional model of FAtiMA is derived from the OCC theory of emotions (Ortony, Clore, and Collins 1988) and the appraisal theory (Smith and Lazarus 1990). Its cognitive layer is based on an efficient implementation of a partial-order-causal-link planning algorithm (Russell and Norvig 2003). A FAtiMA agent, specified through a formal language that is a variant of PDDL2.1 (Fox and Long 2003), is characterised by: (i) a set of internal goals; (ii) a set of action strategies to achieve these goals; and (iii) an affective system. The two main mechanisms controlling a FAtiMA agent are *appraisal* and *coping*. The agent experiences one or more of the 22 emotions of the OCC model based on its appraisal of the current external events against the backdrop of its own goals as well as its subjective tendencies to experience certain emotions instead of others. The agent deals with these emotions by applying problem-focused or emotion-focused coping strategies. When the agent uses a problem-focused coping strategy, it tries to reduce the dissonance between its goals and the external events by acting on the external world to change it. The agent adopts an emotion-focused strategy when it tries to adapt its own emotions to the external events by changing its goals and beliefs based on external circumstances. Both appraisal and coping work at two different levels: *reactive*, which affects the short-term horizon of the agent's behaviour, and *deliberative*, which relates to the agent's long term goal-oriented behaviour. The core of the deliberative layer is a partial-order *continuous* planner that continuously generates plans, triggers the execution of the actions in the plans and monitors all events to detect whether or not the actions under execution are accomplished. FAtiMA interleaves planning and execution so that there is always an appropriate action that the agent can execute. The affective system acts as a powerful heuristic for the planner, because it controls the importance of goals and the selection of actions. The action that triggers the most intense emotion is selected for execution and goals are selected or dropped based on the emotions that they elicit in the character.

Agent's Model and Behaviour: In ECHOES, each learning activity has a FAtiMA agent model associated with it. All these models share the specification of the agent's affective system, so that the agent can maintain the same personality between sessions and establish a trusting relationship with the child. Andy is positive, motivating and supportive. It has a tendency to be happy and does not get frustrated easily. We obtained such behaviour by manipulating its goals and its affective system rules. While Andy's personality does not change between activities, the set of goals that it actively tries to pursue and its action strategies are specified for each learning activity based on: (i) the high-level pedagogical goals of a given activity; and (ii) the specific narrative content of the activity. For example, if the high-level goal of an activity is "Engage in reciprocal interaction" and the content of the activity involves picking flowers in the garden, one of the low-level goals of the agent will be to fill a basket with flowers in collaboration with the child, and its action strategies will demonstrate to the child different ways of engaging in reciprocal interaction, e.g. by choosing between pointing at a flower, looking at it, or saying "Your turn!". Given that ECHOES focuses on supporting joint attention and symbolic use, the agent's actions are either concrete demonstrations of the related skills or actions performed to invite the child to practice those skills (Requirement no. 2). Specifically, SCERTS defines the joint attention and symbolic use as: (i) responding to bids for interaction; (ii) initiating bids for interaction; and (iii) engaging in turn taking. Our agent is able to perform these skills in three different ways: (i) verbally by using simple language or key phrases (e.g., "My turn!" and "Your turn!" for turn-taking); (ii) non-verbally through gaze and gestures, such as pointing at an object from a distance or touching the object; (iii) by combining verbal and non-verbal behaviours. The ECHOES agent is able to make requests, to greet the child by name, to comment on actions or events in the garden and to explore the features of the magic objects populating the garden. This variety of behaviours makes the interaction dynamic enough to keep the child engaged, while retaining a degree of predictability that is essential to supporting the child's attentional focus. Andy always provides the child with positive feedback, especially if the child correctly follows its bids for interaction in task-based activities (Requirement no. 3). If the child does not perform the required action, the agent first waits for the child to act and only after a long pause it intervenes by demonstrating the action and encouraging the child to try again. To provide organisational support, the agent always explains a new activity to the child by using simple language and precise instructions (e.g., "Let's pick all the flowers").

5 The ECHOES Intelligent Engine

Drama Manager: For each session, the drama manager is responsible for establishing the initial state as well as the goals for the agent and for passing this information to the planner. While the overall set of goals to choose from is based on SCERTS, both the initial situation and the specific set of goals for each user in any given session need to be decided based on the user's profile and their interaction history with the system. After delegating the session goals to the agent, the drama manager leaves the agent free to interact with the child without interfering. However, it monitors the unfolding of the interaction and receives input from the user model. If the interaction diverges significantly from the pedagogical goals of the session or the child experiences extreme anxiety or arousal, the drama manager can intervene to keep the interaction on track. For example, it can suspend the execution of the current plan, influence how the planner constructs a new plan, change the overall goals of the session and even drop these goals, if appropriate. Examples of similar drama managers can be found in interactive storytelling systems (e.g. Riedl, Saretto, and Young 2003).

Social-Communication Engine: The role of this engine is to attribute meaning to the child's actions from a social communication standpoint. It builds on low-level information about the actions that the child has performed (e.g. the child has offered an object to the agent) and brings this information to a higher level of abstraction by linking gestures to their social meaning (e.g. the child has responded in an intentional manner to a request from the agent to give it an object). The SCERTS framework, along with the specific context of the interaction between the child and the agent, provide the basis for the set of rules driving this component.

User Model: The user model is intended to estimate the child's cognitive and affective states in real time and feedback this information to other IE components whenever a change in a given state is detected. The user model supports Andy in making informed decisions about how to act appropriately towards the child given the child's estimated intentions, needs and desires (Requirement no. 4). It works based on the real-time information from the touch and eye-gaze systems and produces output at two levels: cognitive and affective. The *cognitive* assessment is facilitated by a *rule-based engine* which estimates the extent to which the child has achieved the goals associated with the session. The rules are based on SCERTS guidelines and precise timing constraints for establishing the child's mastery of joint attention and symbolic use skills. For example, the behaviour "shifts gaze between people and objects" is satisfied if the child shifts gaze spontaneously between a person and an object at least three times and the entire sequence occurs within two seconds. This behaviour must be exhibited at least ten times during a session before the model can infer that the child has mastered it. The *affective* assessment is based on a combination of supervised and unsupervised learning techniques used to estimate the child's level of engagement with the system. Engagement is defined in terms of: *very engaged*, *engaged* and *not engaged*. Engagement is an important indicator of the child's affective state, because it is linked with child's interest and excitement, known to impact learning, while disengagement is linked with boredom and possibly anxiety, known to be detrimental to learning. We assess engagement based on the level of the *flow* experienced by the child (Csikszentmihalyi 1977), where flow is a "mode of experience when the players become absorbed in their activity. This mode is characterised by narrowing of the focus of awareness, so that irrelevant perceptions and thoughts are filtered out, by loss of self-consciousness, by a responsiveness to clear goals and unambiguous feedback, and by a sense of control over the environment". The children are considered "very engaged" when they are fully absorbed by the learning activity proposed by the agent and actively participating in it; "engaged" when they are interested in the current learning activity, but not immersed in it, and they interact with the practitioner in relation to the activity (e.g. child asks questions about the activity); "not engaged" when they do not interact with ECHOES at all, either via gaze or touch. In order to obtain data to train the classifier, we conducted two studies with autistic children using a first prototype of ECHOES and involving a total of 46 children aged 5 to 14 (Porayska-Pomsta et al. 2012). The interactions between the children and the environment were video recorded and then annotated for engagement. We synchronised the video annotations with the system log files, concentrating particularly on how often the child touched the screen. Using these data, we then trained a Support Vector Machine (SVM) classifier using WEKA to predict engagement. Every second, the classifier estimates the subject's level of engagement based on how often they touched the screen in the preceding 1 to 5 seconds. Our tests of engagement estimation based on unsampled data from six children interacting with ECHOES suggest 68% accuracy for a baseline classifier for an overall F-measure of 0.078 using 10-fold cross-validation.

We faced significant challenges in interpreting the child's mental states accurately, which can be largely ascribed to the data being collected *in the wild*, i.e. in real school environments with the children standing and free to move around in unstable lighting conditions. It was particularly difficult to collect reliable data through the eye-tracking systems. In order to deal with user model failures, we supplemented the system with a wizard-of-oz control panel, which allows a human operator to take the role of the drama manager and influence the behaviour of the agent, and used this version of the system for the evaluation studies reported in Section 6. We are currently working on improving the accuracy of the user model based on the data collected during such evaluation studies.

6 Empirical Results

Experimental Design: A large scale multi-site intervention study was conducted to assess the impact of ECHOES and Andy on ASCs children's social communication and their engagement with the environment. The system was deployed in five schools in the UK. Nineteen children with ASCs participated in the study during which they played with ECHOES for ten to twenty minutes, several times a week over an eight week period. To assess each child's initial social communication skills, a structured table-top turn-taking activity (henceforth: "pre") was conducted and their behaviours assessed from video recording of the session. At the end of the intervention, a second table-top session ("post") was conducted to assess generalisation of the social behaviours learned during the use of ECHOES. The SCERTS Assessment Protocol (SAP) (Prizant et al. 2006)

Behaviour	Pre:M(SD)	Post:M(SD)	t(7)	p	sig
Response to social partner	33.7(17.55)	36.37(19.24)	0.219	0.4	no
Initiations to social partner	6.8(4.1))	7.5(9.77)	0.06	0.47	no
Social behaviour	18.3(4.1)	19.5(9.77)	0.07	0.474	no
Sequences of social behaviours	0.4(0.22)	0.125(0.125)	0.798	0.225	no
Speech towards social partner	15.6(8.05)	17.59(17)	0.2	0.42	no
Missed opportunities to respond	26.5(5.7)	17(3.71)	2.57	0.02	yes
Behaviour	**Beg:M(SD)**	**End:M(SD)**	**t(7)**	**p**	**sig**
Response to social partner					
Human	15(10.12)	20(16.63)	1.5713	0.07	no
Andy	8.27(10.12)	10.5(16.63)	1.405	0.096	no
Initiations to social partner					
Human	6.6(7.51)	13(3.26)	1.31	0.089	no
Andy	3.6(6.5)	7.7(4.9)	0.50	0.314	no
Social behaviour					
Human	13.27(7.51)	25(3.26)	3.81	0.002	yes
Andy	19.45(6.5)	31.8(4.9)	1.01	0.19	no
Sequences of social behaviours					
Human	1.72(1.16)	2.3(1.11)	0.43	0.34	no
Andy	2.09(0.79)	3.2(1.98)	0.52	0.3	no
Speech towards social partner	2.72(1.6)	6(2.3)	1.88	0.04	yes
Missed opportunities to respond					
Human	7.5(2.17)	4.7(1.03)	0.4341	0.33	no
Andy	5.45(1.96)	7.1(1.05)	0.52	0.30	no

Table 1: Comparison of manifestation of social behaviours in the children before (Pre) and after (Post) intervention and between the first session with Andy (Beg) and the last (End).

was modified into a finer-grained coding scheme that could be applied to videos of children's interactions with Andy. The modified SAP coding scheme contains sixteen main behavioural categories. Fifteen minute periods during which the children interacted with Andy in the presence of at least one human practitioner were identified for analysis from the beginning ("beg"), middle ("mid"), and end ("end") of the intervention period. Each video was blind-coded by a coder trained in the modified SAP coding scheme. Video annotations were applied by using "ELAN" (a professional tool for complex annotations on video and audio resources) and moderated by a second coder. Due to space limitations, we will not report the results for all the sixteen behavioural categories, but focus on the following six social behaviours that are usually the most severely impaired in children with ASCs: 1. child's response to bids for interaction from a social partner; 2. child's initiation of bids for interaction to a social partner; 3. social behaviour towards social partner and, in particular: (a) using gaze for social referencing (the child looks towards a partner for information) and social sharing (the child initiates joint attention through a combination of gaze and gesture to convey enjoyment and interest); (b) monitoring the attentional focus of the partner; (c) securing the attention of the partner; (d) greeting the partner; and (e) facilitating continuation of turn-taking. 4. Sequences of social behaviours: child's ability to engage in a brief or long reciprocal interaction sequence with a social partner: the child initiates and responds to bids for interaction for at least two (brief) or four (long) consecutive exchanges, where an exchange consists of a turn for the child and a turn for the partner, and at least one exchange is initiated by the child; 5. speech towards agent; and 6. missed opportunities to respond to social partner, where a missed opportunity is a situation where a social partner bids for interaction, but the child does not respond at all, makes a non-contingent response, or initiates a different interaction entirely. Note that both verbal and non verbal social behaviours towards a social partner (the agent or the human practitioner) were coded.

Results concerning Social Behaviours: The graphs in Figure 1 show the total number of times each behaviour was exhibited by the children towards the agent (red line) and the practitioner (blue line) during the pre and post table-top sessions and the the initial and final ECHOES sessions. Table 1 shows two statistical comparisons, one between the pre and post table-top sessions and the other between the first and last ECHOES sessions by reporting means (M), standard deviations (SD), t value (t), degrees of freedom, and significance level (p, sig). As for the pre-post comparison, the results show a positive trend: apart for the sequences of social behaviours, all the other categories register an improvement in the manifestation of the social behaviours between the pre and post table-top test. For example, there was an increase in the number of times the child responded to the practitioner's bids for interaction between the pre (M(SD): 33.7(17.55)) and post (M(SD): 36.37(19.24)) table-top sessions. These improvements reach significance for the "missed opportunity" category. Similarly, the beg-end comparison shows improvements in the social behaviours of the children in regard to both the agent and the practitioner across the ECHOES sessions. For example, there was an increase in the number of child responses to the practitioner (beg (M(SD): 15(10.12)) and end (M(SD): 20(16.63))) as well as an increase in the number of the child initiations to the practitioner (beg (M(SD): 6.6(7.51)) and end (M(SD): 13(3.26))). The improvements reach significance for the "social behaviour" and "speech to agent" categories.

Not many studies on the efficacy of autism intervention, both in human-human and technology-enhanced settings, report statistically significant findings. The heterogeneity in diagnoses of the autistic population makes it difficult to identify group improvements in social behaviours. However, the positive trends identifiable in our experimental results, although not always statistically significant, give us an indication of the potential of ECHOES and offer evidence that several children have benefited from their exposure to ECHOES. Given the increasing complexity of the learning activities and their dependence on turn-taking with Andy across the three sessions, the slight increase in the number of times the children exhibited social behaviours in both comparisons (pre/post and beg/end) is reassuring and may indicate that the child was increasingly regarding Andy as socially credible partner with whom they could interact. Anecdotal evidence supporting this hypothesis comes from a number of children who showed no initial interest in Andy, but spontaneously talked to it and even waved when it walked on the screen in the final sessions. Such behaviours were extremely surprising to teachers and support workers in the school as they believed the children in question to be non-communicative. Similar cases were observed and reported in all the schools that participated in the evaluation.

Results concerning Engagement: We also conducted an analysis on: 1. the overall level of engagement across the three ECHOES sessions; and 2. the level of engagement exhibited by the children in the context of the different learning activities. Note that, based on the definition of the three categories of engagement give in Section 5, children are defined as "engaged" when they interact with both Andy and the hu-

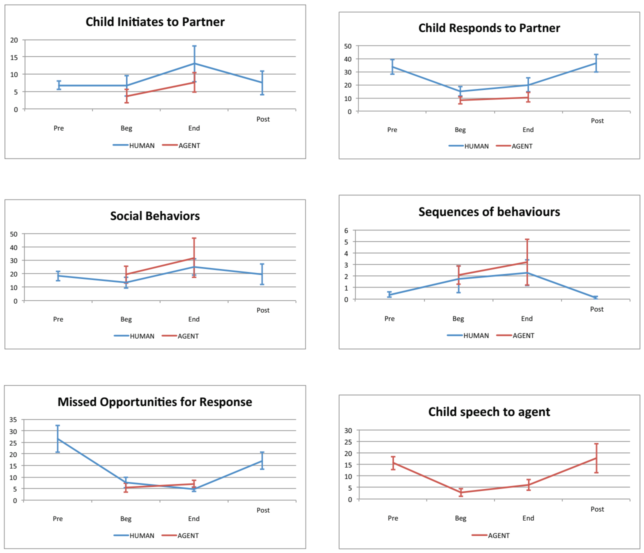

Figure 1: Trends of children's social behaviours before and after intervention and across the ECHOES sessions. Error bars indicate $\pm$ standard error.

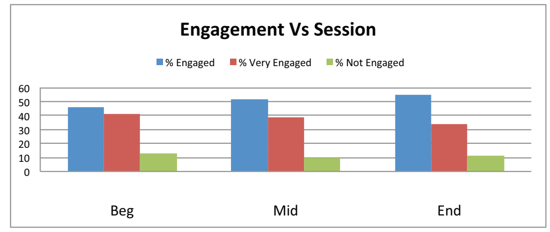

Figure 2: Children's level of engagement across the sessions

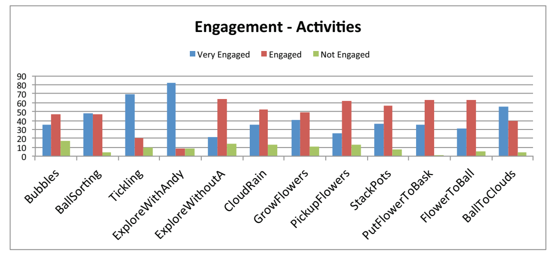

Figure 3: Children's level of engagement in the different ECHOES learning activities

man practitioner and "very engaged" when they interact with Andy only. Therefore, as a measure of the child's progress in social communication, the status of "engaged" is more desirable than "very engaged". Figure 2 shows the children's level of engagement across the three sessions (i.e. "beg", "mid", "end"). The probability of disengagement is negligible; the probability of the child being engaged increases across the sessions, while the probability of the child being very engaged slightly decreases. We consider these trends positive as the children progressively interacted more with the human practitioners and were less absorbed in the system itself. This is a desirable outcome of the intervention: ECHOES was intended to provide ASCs children with a tool for practising social behaviours, as opposed to a game that would reinforce their typical predisposition to social isolation and repetitive and restricted behaviours. Figure 3 shows the children's level of engagement across the 12 learning activities. On average, the children were more "engaged" with the learning activities than "very engaged" or "not engaged", with a higher level of absorption for those activities that were easier to perform (e.g., tickling or exploring the garden). The results suggest that the children interacted with the human practitioners when the tasks were more complicated (e.g., stacking pots or choosing the right flower to pick), probably to ask for support, approval or reassurance.

7 Conclusions

In this paper, we presented our approach to developing, deploying and evaluating an intelligent virtual agent that helps children with ASCs develop social communication skills. Designing a social partner for children, especially autistic children, increases the need for credibility and believability of the agent. We argue that such characteristics come from the ability of the agent to act autonomously and to adapt to individual children in real-time. We believe that the construction of a robust architecture that brings together autonomy, deliberative and emotional reasoning and real-time reactivity continues to be an open problem. Despite the difficulties that we encountered in the development of ECHOES, we managed to produce a prototype that we deployed in-the-wild and the efficacy of which we explored rigorously. To the best of our knowledge, the ECHOES evaluation represents one of the first major evaluations of a pedagogical virtual companion for autistic children conducted in real-school contexts and involving a significant number of children. Although presently we can only report coarse-grained analysis of children's behaviours in relation to Andy, we can already offer evidence that some children have benefited from their exposure to Andy and the ECHOES environment as a whole. Specifically, the experimental results show that, considering the first and the last session with ECHOES, there was an increase in the number of manifestations of almost all the social behaviours, both with respect to the agent and the practitioner, and in some cases this increase reached significance. A possible interpretation of this phenomenon is that Andy's reciprocal interaction with children and its critical role in the learning activities are responsible for children's spontaneous social behaviours. Post-intervention interviews conducted with teachers and carers at the schools participating in the experiments confirm that this is a plausible interpretation.

Acknowledgments

The work presented in this paper has been conducted in the context of two projects: ECHOES (funded by EPSRC and ESRC - Grant: RES-139-25-0395-A) and SHARE-IT (funded by EPSRC - Grant: EP-K012428-1). We are grateful to the ECHOES and SHARE-IT project teams.

References

American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders*. Fourth Edition, Text Revision (DSM-IV-TR).

Aylett, R.; Dias, J.; and Paiva, A. 2006. An affectively-driven planner for synthetic characters. In *Proc. of International Conference in Automated Planning and Scheduling (ICAPS-06)*. AAAI Press.

Bernardini, S., and Porayska-Pomsta, K. 2013. Planning-Based Social Partners for Children with Autism. In *Proc. of the Twenty Third International Conference on Automated Planning and Scheduling (ICAPS-13)*. AAAI Press.

Bernardini, S.; Porayska-Pomsta, K.; Smith, J. T.; and Avramides, K. 2012. Building Autonomous Social Partners for Autistic Children. In *Proc. of the 12th international conference on Intelligent Virtual Agents*, Lecture Notes in Computer Science (LNCS), 46–52. Springer.

Bosseler, A., and Massaro, D. 2003. Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism. *Journal of Autism and Developmental Disorders* 33(6):653–672.

Csikszentmihalyi, M. 1977. *Beyond boredom and anxiety*. San Francisco, CA: Jossey-Bass.

Dias, J., and Paiva, A. 2005. Feeling and reasoning: A computational model for emotional characters. In *Progress in Artificial Intelligence*, volume 3808 of *Lecture Notes in Computer Science*. Springer Berlin, Heidelberg. 127–140.

Fox, M., and Long, D. 2003. PDDL 2.1: An extension to PDDL for expressing temporal planning domains. *Journal of Artificial Intelligence Research* 20:61–124.

Grynszpan, O.; Martin, J.; and Nadel, J. 2008. Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal on Human-Computer Studies* 66(8):628–639.

Milne, M.; Luerssen, M.; Lewis, T.; Leibbrandt, R.; ; and Powers, D. 2010. Development of a virtual agent based social tutor for children with autism spectrum disorders. In *Proc. of the 2010 International Joint Conference on Neural Networks (IJCNN)*, 1–9.

Murray, D. 1997. Autism and information technology: Therapy with computers. In *Autism and Learning: A Guide to Good Practice*. Brookes. 100–117.

Ortony, A.; Clore, G. L.; and Collins, A. 1988. *The Cognitive Structure of Emotions.* Cambridge University Press.

Parsons, S., and Cobb, S. 2011. State-of-the-art of virtual reality technologies for children on the autism spectrum. *European Journal of Special Needs Education* 26(3):355–366.

Porayska-Pomsta, K.; Frauenberger, C.; Pain, H.; Rajendran, G.; Smith, T.; Menzies, R.; Foster, M. E.; Alcorn, A.; Wass, S.; Bernardini, S.; Avramides, K.; Keay-Bright, W.; Chen, J.; Waller, A.; Guldberg, K.; Good, J.; and Lemon, O. 2012. Developing technology for autism: An interdisciplinary approach. *Personal and Ubiquitous Computing* 16(2):117–127.

Porayska-Pomsta, K.; Anderson, K.; Bernardini, S.; Guldberg, K.; Smith, T.; Kossivaki, L.; Hodgings, S.; and Lowe, I. 2013. Building intelligent authorable serious game for autistic children and their carers. In *Proceedings of the 10th International Conference on Advances in Computer Entertainment (ACE 2013)*.

Prizant, B.; Wetherby, A.; Rubin, E.; and Laurent, A. 2003. The SCERTS model: A Transactional, Family-Centered Approach to Enhancing Communication and Socioemotional Ability in Children with Autism Spectrum Disorder. *Infants and Young Children* 16(4):296–316.

Prizant, B.; Wetherby, A.; Rubin, E.; Laurent, A.; and Rydell, P. 2006. *The SCERTS® Model: A Comprehensive Educational Approach for Children with Autism Spectrum Disorders*. Brookes.

Putnam, C., and Chong, L. 2008. Software and technologies designed for people with autism: what do users want? In *10th International ACM SIGACCESS conference on computers and accessibility*, 3–8.

Riedl, M.; Saretto, C. J.; and Young, R. M. 2003. Managing interaction between users and agents in a multi-agent storytelling environment. In *Proceedings of the second international joint conference on autonomous agents and multiagent systems (AAMAS-03)*, 741–748. New York: ACM.

Russell, S., and Norvig, P. 2003. *Artificial Intelligence: A Modern Approach.* Prentice Hall. Second Edition.

Smith, C. A., and Lazarus, R. S. 1990. Emotion and adaptation. In Vlahavas, I., and Vrakas, D., eds., *Handbook of Personality: Theory and Research*. New York: Guilford: L.A. Pervin (Ed.). chapter 23, 609–637.

Tartaro, A., and Cassell, J. 2008. Playing with virtual peers: bootstrapping contingent discourse in children with autism. In *Proc. of the 8th International Conference for the Learning Sciences (ICLS-08)*, 382–389.