GAIPS tech. rep. series, n. GAIPS-TR-001-19, March 2019

INTELLIGENT AGENTS AND SYNTHETIC CHARACTERS GROUP

An Approach for Personalized Social Interactions Between a Therapeutic Robot and Children with Autism Spectrum Disorder

Kim Baraka
INESC-ID / Instituto Superior Técnico, Universidade de Lisboa
Porto Salvo, Portugal
Robotics Institute, Carnegie Mellon University
Pittsburgh, PA, USA

Marta Couto
Hospital Garcia de Orta, EPE.
Almada, Portugal

Francisco S. Melo, Ana Paiva
Intelligent Agents and Synthetic Characters Group
INESC-ID, Lisbon
Portugal

Manuela Veloso
School of Computer Science, Carnegie Mellon University[a]
Pittsburgh, PA, USA

For high-variability populations, such as individuals with Autism Spectrum Disorders (ASD), the importance of personalization and adaptation mechanisms in Human-Robot Interaction becomes crucial. This technical report presents an algorithmic method for personalization of robotic behavior in structured social interactions with ASD children, based on state-of-the-art diagnostic models. In a first step, we leverage the structure of the diagnostic procedure to build robotic behaviors on a NAO humanoid robot, aimed at eliciting target behaviors from the child. Through appropriate sequencing of possible actions, the robot is able to assess a child's behavioral profile and use it to personalize the interaction. To test our method, we developed a semi-autonomous robotic scenario where a humanoid robot interacts with a child with ASD through interactive storytelling, focusing on social prompts related to deficits in attention, one of the core impairments of ASD. We present the design and methodology of an evaluation study run with 11 young ASD children in a child development center in Portugal.

[a]Manuela Veloso is currently the Head of JPMorgan Artificial Intelligence

Contents

GAIPS/INESC-ID
TagusPark, Edifício IST
Av. Prof. Dr. Cavaco Silva
2780-990 Porto Salvo
Portugal

Tel.: +351 214 233 508
Fax: +351 214 233 290
http://gaips.inesc-id.pt/

Corresponding author:
Kim Baraka
E-mail: kbaraka@andrew.cmu.edu
https://www.kimbaraka.com/

1 Introduction

Children with Autism Spectrum Disorders (ASD) are a prime example of a high variability population, both across and within individuals. This is specially true concerning their behavioral profiles, where we see immense variability along a set of different scales [1]. In recent years, there has been a strong interest in using robotic technologies to assist such individuals in a variety of therapeutic tasks [11] [4] [10]. Robots possess many desirable features that make them attractive both for individuals affected by ASD, who seem to show general affinity with technology, as well as for people studying autism. They allow for more effective and objective data gathering, and offer more repeatability, control, and potentially flexibility [3]. If repeatability and control have been the main focus of existing socially assistive robots, flexibility, i.e. the ability to accommodate for inter- and intra- individual differences, remains a challenging but necessary subject to tackle.

This paper presents a method for personalization of structured robotic behaviors in the context of a social interaction. The same way diagnosis data can inform the personalization strategies of therapists in their interactions with patients, we believe that a robot could build a useful behavioral profile of the patient through interactions similar to those used in diagnosis, allowing it to personalize and adapt to the abilities and needs of a diverse range of patients. Personalization and adaptation are widely used strategies amongst autism therapists, but the challenges of achieving basic such mechanisms in robot-assisted therapeutic tasks are numerous. They include:

(1) Profiling Building useful profiles of children interacting with robots consists in assessing features related to their interaction with the robot. Profiling is a challenging task because: (a) Child response to robots may significantly differ from response to humans, which means there might not be a systematic way to predict response with a robot given data on response to a human therapist. (b) The cost of exploration may be high. Individuals with ASD are often extremely sensitive, and a single 'wrong step' in the robot's behavior may result in dramatic consequences. (c) The amount of data that a robot can collect with a specific child is limited, which poses the challenges of estimating child features that are useful for the interaction from scarce data.
In our work, we base our feature assessment method on standard diagnostic procedures widely used by human therapists.

(2) Personalization Adapting robot behavior according to each profile is a separate research question, which requires domain knowledge. What strategy works best for each profile? How can its efficacy be measured?
In this work, our personalization strategy aligns with typical strategies followed by human therapists that have been shown to promote learning in the long-term.

(3) Integration in naturalistic context Since most ASD therapy tasks rely on aspects of social interaction, it is necessary that they be integrated in an engaging scenario with a specific meaning and progression. Maintaining stable engagement levels with such a population is particularly challenging and also particularly helpful as it reduces uncertainty in the robot's ability to predict children's response.
For evaluation purposes, we integrate tasks that allow for personalization of robot behavior within a larger interactive storytelling scenario.

Personalization plays a crucial role in the context of autism therapy, which focuses in large part on training children to better understand and respond to a range social cues. The most effective approach on the long run is to tailor the 'just right challenge' to each individual [12]. Effectively, this strategy translates into finding the right balance between making social cues 'easy enough' to limit task duration and maintain engagement, but 'hard enough' to promote improvement over time by challenging current child abilities.

Along those lines, our approach starts by leveraging the structure of diagnostic interactions to model the interaction between a child with ASD and another agent (therapist or robot) in two structured tasks. Inspired by procedures from the Autism Diagnosis Observation Schedule (ADOS-2) [6], a state-of-the-art diagnostic tool, we develop a set of prompting behaviors on a NAO humanoid robot, aiming at eliciting a target response from the child in two attention-related tasks involving screens. The robot behaviors, which we call 'presses' in accordance with the ADOS-2 terminology, fall under a hierarchy organized along a scale of levels of explicitness, adapted to a range of possible child behavioral profiles. Based on these robotic presses, we develop a control architecture that allows the robot to prompt the child in different sequences of presses according to its mode of operation - Assess, Explore or Exploit - in the context of a social interaction scenario.

To test the validity of our approach, we devise an interactive storytelling scenario involving a NAO humanoid robot and two controllable screens showing cartoon excerpts. In our scenario, the robot is able to estimate the behavioral profile of the child with sparse measurements, utilize it to personalize and promote learning, as well as allowing for potential exploration of alternative policies, in a manner that is embedded within the scenario. We present the details of an evaluation study we ran with 11 children with ASD in a child development center in Portugal.

2 A diagnosis-inspired robotic prompting scheme

In this section, we describe our robotic prompting scheme developed for a NAO humanoid robot, and inspired by the 'algorithmic' nature of two ADOS-2 activities, related to joint attention and reponse to name. After describing the interaction setup considered, we explain how we developed robotic presses inspired by ADOS-2. We then discuss our flexible robot control architecture, that allows for three different modes of operation: Assess, Explore, Exploit.

2.1 Interaction setup

The physical setup we consider, shown in Fig. 1, was inspired by the work of Warren et al. [13], who demonstrated its suitability for young ASD children. We found this scenario to be attractive to explore the idea of personalization of attention-related interactions, because it allows for both control and flexibility, as compared to scenarios involving physical objects, portable digital devices (e.g., tablets) [2], or scenarios where the child moves around the space [7]. The setup consists of a NAO robot standing on a table, at which the child is seated, and two 49.4 cm LCD screens positioned at around a 90 degree angle on both sides of the child's chair.

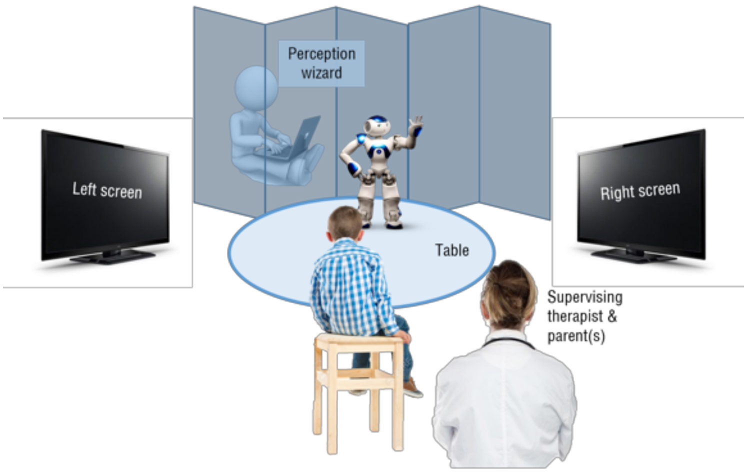

Figure 1: The interaction setup considered in this work. Figure only for illustrative purposes; relative positions and sizes of the components is not exact.

The robot can engage in two tasks, whose goals are the following:

- **Joint Attention (JATT) task**: Direct the child's gaze from the robot to a target screen where a video will play.
- **Name Calling (NAME) task**: Direct the child's gaze from the video on the screen back to the robot.

A 'perception Wizard', whose main role is to provide the robot with information about the gaze response of the child through a computer interface, hides behind a single-sided mirror at an angle that maximizes the view to the scene. During each of the two tasks, he/she triggers a 'hit' event whenever the child performs the target behavior for that task. While eye-tracking or head-tracking technology were available for us to use, we decided to rely on human perception, as such technologies are either too invasive (a problem for ASD children with sensory overload) or too inaccurate, especially for children with attention impairments who tend to move considerably. For the JATT task, a hit triggers a short video snippet. For the NAME task, a hit stops the video playing on the screen where the child is looking.

Screens are individually controllable, through a single processing unit. The 'Wizard' 's computer runs the main software to control both the robot and the screens. A wired network connection through a switch between all computing units is used to minimize delays, which are crucial for real-time applications that require low latency like ours. We used the Thalamus framework [8] to facilitate communication between the distributed modules.

For safety, the robot's feet were stuck to the table using tape to avoid robot falls, as we have noticed that some children are particularly keen on touching and poking the robot.

2.2 Developing a hierarchy of robotic presses inspired by ADOS-2

As part of the ADOS-2 diagnostic tool, there exist two systematic 'algorithms' for evaluating response to joint attention and response to name, through a hierarchy of presses with increasing level of explicitness. Each press corresponds to a more or less explicit action taken by the therapist with the same aim of elicit a target behavior on the child's part. The ADOS-2 presses and the target child behaviors are summarized in columns 2-4 of Table 1.

Inspired by the structure of the ADOS-2, we developed a similar hierarchy of presses for a NAO humanoid robot, aiming to elicit a target behavior from the child. Column 5 of Table 1 summarizes our developed robotic presses. Note that the aim was not to replicate with high fidelity the content of the ADOS-2 presses, but rather to come up with a similar hierarchy that is adapted to our scenario, and accounts for a range of responses across the hierarchy. Also, to ensure an increasing level of explicitness for the presses, we structured them such that press $i+1$ is a replica of press i with an added element that either adds intensity to the stimulus (e.g., sound on top of video) or facilitates the understanding of the press (e.g., pointing added to gaze). We used the SERA software architecture [9] to control the robot's behaviors. The speech is automatically generated by the NAO built-in Text-to-Speech (TTS) engine.

We finetuned our presses based on pilot trials with 4 typically developing (TD), 2 ASD, and 1 minimally ASD children. Specifically, for task JATT, special care had to be taken with the behavior of the screens, as it seemed from our pilots that the sharp transition from a black screen to an image or video was a very salient stimulus that transiently overpowered the robot's role. For this reason, we decided to pre-load a static picture on both screens, corresponding to the first frame of the video to be shown, and to keep the brightness of the screens on a low setting.

2.3 Child behavioral profile

In ADOS-2, the therapist goes through the presses sequentially from least to most explicit until the target response is elicited, and records the level of the first successful press. This number can be seen as a measure of abnormality of response to the task. In this work, since we consider two tasks, the child behavioral profile is represented as a pair of features $(f_{\mathrm{JATT}}, f_{\mathrm{NAME}})$, where f is the lowest press level at which a hit is observed. If none of the 4 press levels cause a hit to occur, we assign to f a value of 5. In a typical ADOS-2 session, f_{JATT} and f_{NAME} are measured only once. In a robotic scenario however, we expect much greater variability in the response due to the novelty effect associated with the robot, as well as the scenario as a whole. For this reason, estimating values of f accurately may require several samples. Given N measurements $f_1, f_2 \ldots, f_N$, we estimate f as:

$$\hat{f} = \begin{cases} \mathrm{rnd}(\sum_1^N \frac{f_i}{N}) & \text{if } \sum_1^N \frac{f_i}{N} \mod 2 \neq 0.5 \\ \mathrm{rnd}(\sum_2^N \frac{f_i}{N-1}) & \text{if } \sum_1^N \frac{f_i}{N} \mod 2 = 0.5, \end{cases} \tag{1}$$

Table 1: Summary of our robotic presses with increasing levels of explicitness (1-4), inspired by the hierarchical structure of ADOS-2

Task	Target behavior	Press	ADOS-2 presses	Our robotic presses
JATT	Look at target object (screen)	p_1	Gaze shift from child to target object + "[Name], look!"	**(Gaze + speech)** Gaze shift from child to target screen + "[Name], look!" (Static picture on both screens)
		p_2	Gaze shift + "[Name], look at that!"	**(Gaze + speech + pointing)** Gaze shift + "[Name], look at that!" + pointing (Static picture on both screens)
		p_3	Gaze shift + "[Name], look at that!" + pointing	**(Gaze + speech + pointing + video)** Gaze shift + "[Name], look at that!" + pointing + muted video on target screen (Static picture on other screen)
		p_4	Activate the target object (toy)	**(Gaze + pointing + video + sound)** Gaze shift + "[Name], look at that!" + pointing + video with localized sound on target screen (Static picture on other screen)
NAME	Look at person/robot	p_1	Call name	"[Name!]"
		p_2	Ask the parent/caregiver to call name	"[Name], look over here!".
		p_3	Ask the parent/caregiver to make a familiar sound	"[Name], look over here!" + Blinking lights
		p_4	Ask the parent/caregiver to do whatever necessary to get the child's attention	"[Name], look over here!" + Blinking lights + Waving arm

where rnd() is the rounding to the nearest integer operation. In other words, in case of an estimate lying exactly in the middle of two levels, we omit the first sample, given that it is more prone to novelty factors and is hence, in comparison to more recent samples, less reflective of subsequent performance of the child on the task. Equation 1 applies for estimating both f_{JATT} and f_{NAME}.

Examples ($N = 4$):
$f_1 = 3, f_2 = 3, f_3 = 4, f_4 = 2 \rightarrow \hat{f} = \text{rnd}(\frac{3+3+4+2}{4}) = 3$
$f_1 = 3, f_2 = 3, f_3 = 2, f_4 = 2 \rightarrow \hat{f} = \text{rnd}(\frac{3+2+2}{3}) = 2$

2.4 Prompting logic

The ADOS-2 'algorithm' requires the therapist to go over the presses by increasing level of explicitness until a hit is observed. Outside the context of assessment however, one may want to consider other sequences of presses than the least-to-most explicit ones. In particular, if a behavioral profile of the child $(f_{\text{JATT}}, f_{\text{NAME}})$ is available, it is natural to start with press $p_{f_{\text{JATT}}}$, since we know that lower level presses have a low probability of hit.

We call a *prompting sequence* a sequence of presses to be performed by the robot in a sequential way until hit occurs or the sequence is exhausted. In this work, we restrict the prompting sequence length to 4, for consistency with the maximum number of steps in the ADOS-2 algorithms. Our robot control architecture is general enough to allow for arbitrary prompting sequences to be followed.

Before starting the execution of the task, the robot first generates a prompting sequence, i.e., a plan of sequential actions to be taken, through a prompting sequence generation module, which receives some parameters from a high-level decision maker. A prompting sequence execution module executes the presses on the robot sequentially, until either a hit is triggered by the 'Wizard' or the sequence is exhausted. The trigger of the next press in the sequence is a timeout in case no hit occurs. Based on our pilots, we set the duration of the timeout to 3.5 seconds. Fig. 2 shows the relation between the different modules of the robot control architecture.

2.5 Prompting modes

As can be seen in the upper left part of Fig. 2, our architecture allows for three prompting modes: *Assess*, *Exploit*, and *Explore*, which effectively translate into different prompting sequence generator outputs.

- ***Assess*** *mode*: The robot follows the ADOS-2 hierarchical 'algorithm', typically used by psychologists for assessment. The prompting sequences for this mode are all of the form: p_1, p_2, p_3, p_4. This mode enables the robot to build a behavior profile of the child by recording the least explicit press level at which the child responds for the two tasks.

- ***Exploit*** *mode*: The robot personalizes the interaction strategy according to a given child profile $(f_{\text{JATT}}, f_{\text{NAME}})$. In this mode, for a given task, the first two presses in the prompting sequence are repetitions of press p_f. The last two

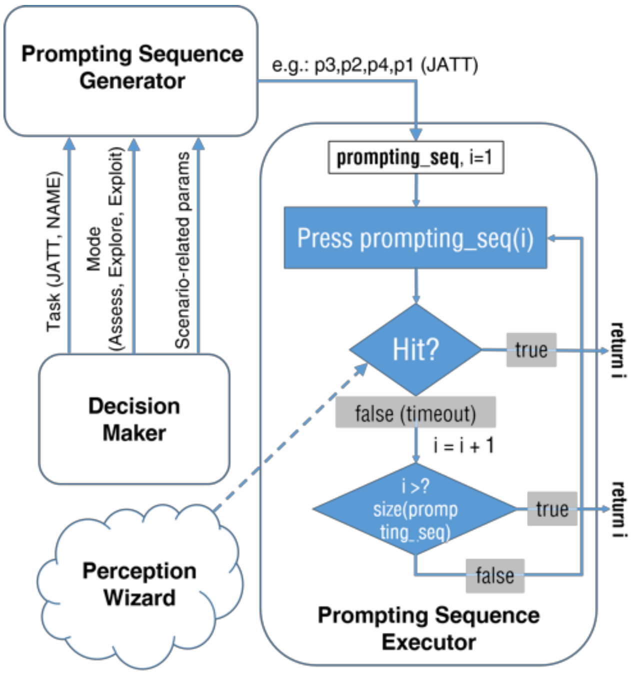

Figure 2: Robot prompting control architecture

presses are repetitions of press p_{f+1}. In the edge cases where $f = 4$ or $f = 5$, this mode generates 4 repetitions of p_4. Such sequences, utilizing knowledge encoded in the child profile, are *personalized* while also being *adaptive*, as the robot increases the explicitness of the press if no hit is observed within two repetitions of the personalized press.

Examples:
$f = 2 \rightarrow$ Prompting sequence: p_2, p_2, p_3, p_3
$f = 3 \rightarrow$ Prompting sequence: p_3, p_3, p_4, p_4
$f = 4 \rightarrow$ Prompting sequence: p_4, p_4, p_4, p_4.

It is important to mention that the goal of this mode is not to minimize the number of presses needed to observe a hit, otherwise the robot could always select the most explicit press p_4. Instead, in alignment with therapeutic goals [5], this mode chooses the least explicit press that has been shown to work on a particular child, in order to promote learning (in the long term) without compromising too much on task performance (expected number of repetitions needed for a hit to occur).

- ***Explore*** *mode*: In this mode, the robot samples the presses of the prompting sequence in a uniformly and independently random fashion.

Note that in any of the modes presented above, the prompting sequence represents a plan, whose execution may be aborted if a hit occurs, i.e. if the child

performs the target behavior. Note that our robot control architecture allows for more modes than the ones above, however those were the three modes that fit our scenario and goal in this work.

In our HRI scenario, in a first phase, we set the robot to the Assess mode, and let it collect samples to estimate the child's behavioral profile. In a second phase, the robot alternates between the Explore and Exploit mode. To reduce the complexity of the data collected, the Explore mode does not affect the estimated child profile used by the Exploit mode.

3 Child-Robot Interaction Scenario

In order to test our robotic prompting scheme in the context of an extended social interaction, we implemented an interactive storytelling scenario, where short excerpts of an animated cartoon on the screens regularly support and illustrate the robot's speech delivery. The JATT task is repeatedly used throughout the interaction to direct the child's attention to one of the two screens where the cartoon excerpt is to be shown. Following this task, the robot uses the NAME task to call the child's attention back to it and resume the storytelling.

3.1 Story design

The story we chose is based on an episode of a Japanese cartoon, Ox Tales, dubbed in European Portuguese. Popular in the previous generation, this amusing cartoon is much lesser known by the younger generation, which reduces the chances of current children having strong (positive or negative) feelings about it. The episode was selected based on the simplicity of the plot and the presence of simple actions for the child to imitate, which the robot uses to engage the child throughout the story. We transcribed, simplified and rewrote the video episode in a storytelling style with simple language to ensure that children with different linguistic levels would be able to follow the story. We then edited and adapted the length and organization of the story based on our pilot trials, aiming at optimizing for child engagement, clarity of robot speech, and plot simplicity.

In parallel to the verbal content of the story, we extracted and edited 12 snippets of 12 seconds each from the cartoon that showed interesting actions throughout the story, including 4 whose aim is to introduce a specific character.

3.2 Robot behavior during storytelling

The robot used NAO's built-in European Portuguese TTS engine for both the storytelling part and the interactive tasks. Even though pre-recorded voice can be much more engaging and natural-feeling for storytelling, the choice of TTS aligned with our long-term goal of a personalized and adaptive solution that includes modulating speech content, and as a result we opted for the greatest level of reliable autonomy possible on the robot side.

To increase the expressivity of the robot during storytelling, we animated it with a 'breathing behavior' consisting of swinging its weight back and forth between each

leg at a rate of 30 times per min. We also added expressive hand gestures, randomly alternating between left and right, inspired by simple gestures typically used by storytellers.

3.3 Interaction timeline

The interactive storytelling scenario alternates between storytelling and interactive prompting as described in the previous sections. Our scenario consists of two consecutive phases: an assessment phase in which the robot presents the characters of the story and a main interaction phase consisting of the actual storytelling. In both phases, the robot uses the cartoon snippets in the prompting tasks. We tried to balance the number of words as much as possible between the different story parts defined by the occurrence of the tasks. Any hit or timeout in the JATT task triggers the 12-second video snippet of the corresponding part of the story. Any hit in the NAME task turns both screens to black for a short period of time, then updates both screens with a new static image corresponding to the next part of the story. To increase the children's engagement, throughout the story, we relied on questions such has "What do you think will happen?", as well as moments where the robot prompted the child to imitate a a total of 4 movements related to the plot. Between the two phases and at the end of the story, the cartoon theme is played on both screens. Fig. 3 shows the timeline of the interaction.

In the assessment phase, the robot is in Assess mode, and performs each task a total of 4 times. It uses the recorded levels of response to estimate the behavioral profile of the child. In the main interaction phase, the robot alternates between the Exploit and Explore modes, performing each task a total of 8 times. The choice of alternating between the two modes served our experimental purposes, to gather data in all three operational modes of the prompting algorithms introduced. Finally, we remind the reader that the Exploit mode only relies on the result from the assessment phase, and unlike some existing machine learning algorithms which interleave Exploration and Exploitation in their policies, is not influenced by the results of the Explore mode.

4 Evaluation study

We evaluated our approach with young children with ASD, using the scenario described in the previous section. This section provides details about this evaluation study.

4.1 Participants

We recruited 11 children with ASD from the Child Development Center at the Hospital Garcia de Orta in Alamda, Portugal, to participate in the study. The criteria for selection were: between 2 and 6 years old, and ASD diagnosis between mild and severe. In addition to these criteria, we consulted with the therapist working with those children asking whether they thought the child would respond well to this type of scenario (sitting on a chair for a relatively long period of time),

- **Baseline phase**

  <Introduce character 1>

  JATT task (Show video of the character)

  NAME task

  Record $\hat{f}_{JATT,1}$ *and* $\hat{f}_{NAME,1}$

  Repeat above for characters 2,3 and 4 and

  record $\hat{f}_{JATT,2}, \dots, \hat{f}_{JATT,4}$ *and* $\hat{f}_{NAME,2}, \dots, \hat{f}_{NAME,4}$

  Compute child behavioral profile $(\hat{f}_{JATT}, \hat{f}_{NAME})$ *using equation (1)*

- **Main interaction phase**

  <Start story>

  JATT Task - *Mode* ***Exploit***

  NAME Task - *Mode* ***Exploit***

  <Continue story>

  JATT Task - *Mode* ***Explore***

  NAME Task - *Mode* ***Explore***

  <Continue story>

  ...

  JATT Task - *Mode* ***Exploit***

  NAME Task - *Mode* ***Exploit***

  <Continue story>

  JATT Task - *Mode* ***Explore***

  NAME Task - *Mode* ***Explore***

  <Continue and end story>

Figure 3: Chronological timeline showing the high-level algorithm followed by the robot throughout the scenario. In the main interaction phase, the robot performs both tasks a total of 8 times.

or if there were any factor that may not make them suitable for our scenario (e.g., fear of robots). The ages of our sample ranged between 2 years 9 months and 7 years 1 month ($M = 4.64$, $SD = 1.36$). Seven were male (63.6%) and four female (36.4%). Three children had low severity scores, 6 moderate and 2 severe. Three of the participants (27.3%) had interacted with a robot before (but not NAO) in the context of a separate study. All participants successfully completed the session, except for one who only completed the first phase.

4.2 Experimental procedure

After obtaining informed consent from the child's parent, the experimenter brought the child into the experiment room, along with their parent(s)/caregiver(s). Before initiating the session, the child was given time to explore the robot, encouraged to touch it and talk to it. During this initial time, the robot was controlled progressively

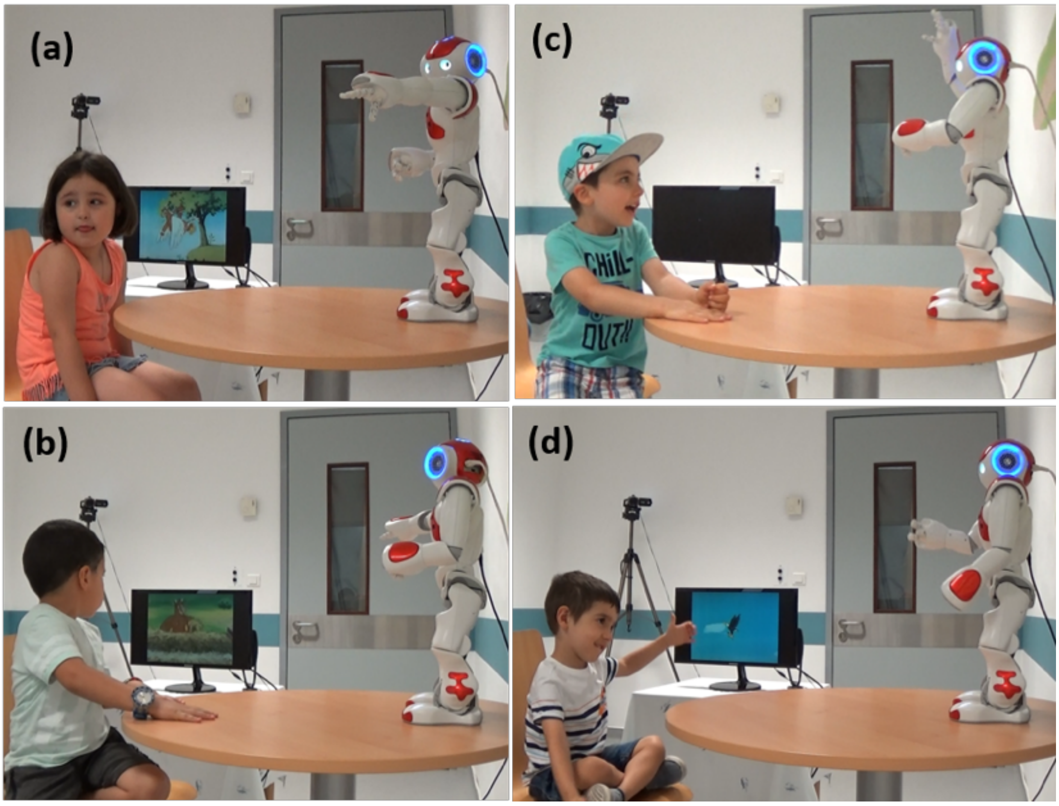

Figure 4: Snapshots from the experimental sessions showing JATT hits for the right (a) and left (b) screen (not visible), a NAME hit for p_4 (c), and a child imitating the robot's movement as instructed during storytelling for increased engagement.

by the 'Wizard' using a library of possible actions to attract the child in case of lack of interest, or to calm the child in case of fear or distress. After the child is seated and ready to interact with the robot, the semi-autonomous control of the robot is initiated. The robot starts by introducing itself and asks for the child's name until the child responds (or the parent, in case of failure). From there on, the experiment timeline outlined in Fig. 3 starts. The total session time ranged between 15 and 20 minutes approximately. Fig. 4 shows some snapshots from different sessions.

The parents were instructed to minimally intervene, especially during the robot prompting tasks, so as not to bias our results. During the robot's prompting tasks, the experimenter followed strict guidelines when intervention was needed. She only intervened to make sure the child was looking at the robot before the robot initiated the JATT task, and at the screen (or at least away from the robot) for the NAME task, both of which are important pre-conditions for the tasks we are studying.

The role of the 'Wizard' was played by a second experimenter during all the sessions. To ensure that there was no bias in the data he provided, we asked an autism therapist, who was agnostic to the aims of the study, to separately record her coding of children's responses for later comparison. Since she was not familiar enough with the 'Wizard' interface, we decided that it was best for her not to operate

the interface directly, as a low latency was crucial when triggering hits.

In the assessment phase, the choice of screen (left/right) in the JATT tasks alternated between consecutive tasks, and the choice of first screen was counterbalanced across participants. In the main interaction phase, the choice of screen was randomly selected while ensuring equal left/right proportions for each participant and not allowing more than two consecutive instances on the same screen, in order to minimize any learning effect. The choice of the firstly selected prompting mode (Exploit/Explore) was also counterbalanced.

The data collected consists of the robot logs, recording the presses selected by the robot and whether they resulted in a hit or a miss, as well as videos of the interactions. We plan to analyze this data to evaluate the effectiveness of our personalization approach, as well as to inform further personalization mechanisms.

5 Conclusion

This technical report outlined the details of a robotic prompting system aimed at therapeutic tasks with children with ASD, and inspired by the ADOS-2 diagnosis tool. We ran an evaluation study with 11 children at a child development center in Portugal. The data collected in the study will be used to evaluate the effectiveness of the personalization approach, as well as to build models of children response in the tasks 6we consider, in order to create better personalization strategies based on model-based optimization techniques.

References

[1] American Psychiatric Association et al. *Diagnostic and statistical manual of mental disorders (DSM-5®).* American Psychiatric Pub, 2013.

[2] Beatriz Bernardo, Patrícia Alves-Oliveira, Maria Graça Santos, Francisco S Melo, and Ana Paiva. An interactive tangram game for children with autism. In *International Conference on Intelligent Virtual Agents*, pages 500–504. Springer, 2016.

[3] John-John Cabibihan, Hifza Javed, Marcelo Ang, and Sharifah Mariam Aljunied. Why robots? a survey on the roles and benefits of social robots in the therapy of children with autism. *International journal of social robotics*, 5(4): 593–618, 2013.

[4] David Feil-Seifer and Maja J Matarić. Socially assistive robotics. *IEEE Robotics & Automation Magazine*, 18(1):24–31, 2011.

[5] Jillian Greczek, Edward Kaszubski, Amin Atrash, and Maja Matarić. Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pages 561–566. IEEE, 2014.

[6] Catherine Lord, Michael Rutter, Pamela Dilavore, Susan Risi, Katherine Gotham, and Somer Bishop. *Autism diagnostic observation schedule, second edition.* Western Psychological Services, CA, 2012.

[7] Francisco S. Melo, Alberto Sardinha, David Belo, Marta Couto, Miguel Faria, Anabela Farias, Hugo Gambôa, Cátia Jesus, Mithun Kinarullathil, Pedro Lima, Luís Luz, André Mateus, Isabel Melo, Plinio Moreno, Daniel Osório, Ana Paiva, Jhielson Pimentel, Jo ao Rodrigues, Pedro Sequeira, Rubén Solera-Ure na, Miguel Vasco, Manuela Veloso, and Rodrigo Ventura. Project INSIDE: Towards Autonomous Semi-Unstructured Human-Robot Social Interaction in Autism Therapy. *Artificial Intelligence in Medicine*, 2019 (to appear).

[8] Tiago Ribeiro, Marco Vala, and Ana Paiva. Thalamus: Closing the mind-body loop in interactive embodied characters. In *International Conference on Intelligent Virtual Agents*, pages 189–195. Springer, 2012.

[9] Tiago Ribeiro, André Pereira, Eugenio Di Tullio, and Ana Paiva. The sera ecosystem: Socially expressive robotics architecture for autonomous human-robot interaction. In *AAAI Spring Symposium Series*, 2016.

[10] Ben Robins, Kerstin Dautenhahn, R Te Boekhorst, and Aude Billard. Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? *Universal Access in the Information Society*, 4(2):105–120, 2005.

[11] Brian Scassellati, Henny Admoni, and Maja Matarić. Robots for use in autism research. *Annual review of biomedical engineering*, 14:275–294, 2012.

[12] Roseann Cianciulli Schaaf and Susanne Smith Roley. *Sensory integration: Applying clinical reasoning to practice with diverse populations.* PRO-ED, Incorporated, 2006.

[13] Zachary E. Warren, Zhi Zheng, Amy R. Swanson, Esubalew T. Bekele, Lian Zhang, Julie A. Crittendon, Amy F. Weitlauf, and Nilanjan Sarkar. Can Robotic Interaction Improve Joint Attention Skills? *Journal of Autism and Developmental Disorders*, 45(11), 2015. ISSN 15733432. doi: 10.1007/s10803-013-1918-4.

Acknowledgements

This research is partially supported by the CMUPERI/HCI/0051/2013 grant, associated with the CMU-Portugal INSIDE project (http://www.project-inside.pt/), as well as national funds through Funda cão para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2019. The views and conclusions contained in this document are those of the authors only.

GAIPS Technical Report Series

This report is part of the GAIPS Technical Report Series. The reports currently available are:

[a] P. Sequeira, F.S. Melo, A. Paiva. Learning by appraising: An emotion-based approach for intrinsic reward design Tech. Rep. GAIPS-TR-001-12, GAIPS/INESC-ID, March 2012.

[b] P. Sequeira, F.S. Melo, A. Paiva. Associative metric for learning in factored MDPs based on classical conditioning. Tech. Rep. GAIPS-TR-002-12, GAIPS/INESC-ID, June 2012.

[c] S. Mascarenhas, R. Prada, A. Paiva. Social importance dynamics: A model for culturally-adaptive agents. Tech. Rep. GAIPS-TR-001-13, GAIPS/INESC-ID, April 2013.

[d] S. Mascarenhas, R. Prada, A. Paiva. Planning culturally-appropriate conversations using adjacency pairs. Tech. Rep. GAIPS-TR-002-13, GAIPS/INESC-ID, April 2013.

[e] P. Sequeira, F.S. Melo, A. Paiva. Emergence of emotion-like signals in learning agents. Tech. Rep. GAIPS-TR-003-13, GAIPS/INESC-ID, October 2013.

[f] S. Mascarenhas, S. Rodrigues, J. Dias, A. Paiva. A process model of empathy for virtual agents. Tech. Rep. GAIPS-TR-004-13, GAIPS/INESC-ID, November 2013.

[g] B. Bernardo. An Interactive Tangram Game For Autistic Children Tech. Rep. GAIPS-TR-001-16, GAIPS/INESC-ID, November 2016.

[h] F. Correia. EMYS: A Social Robot that Plays "Sueca" Tech. Rep. GAIPS-TR-002-16, GAIPS/INESC-ID, November 2016.

[i] J. Campos, J. Guerreiro, B. Ravenet, R. Prada, A. Paiva. Computer supported training of joint investigation teams Tech. Rep. GAIPS-TR-001-17, GAIPS/INESC-ID, May 2017.

[j] J. Campos, J. Guerreiro, B. Ravenet, R. Prada, A. Paiva, E. Jaspaert, G. Vervaeke, G. Nahari, A. Roth, S. Kraus. Designing a training tool for transnational joint investigation teams Tech. Rep. GAIPS-TR-002-17, GAIPS/INESC-ID, May 2017.

[k] F.S. Melo, A. Sardinha, M. Faria, A. Paiva, R. Solera, M. Vasco, M. Kinarullathil, P. Lima, L. Luz, A. Mateus, J. Pimentel, R. Ventura, M. Couto, A. Farias, I. Melo, H. Gambôa, D. Belo, D. Osório, J. Rodrigues, P. Sequeira, M. Veloso Project INSIDE: Towards Autonomous Semi-Unstructured Human-Robot Social Interaction in Autism Therapy Tech. Rep. GAIPS-TR-001-18, GAIPS/INESC-ID, March 2018.

[l] K. Baraka, M. Couto, F.S. Melo, A. Paiva, M. Veloso An approach for personalized social interactions between a therapeutic robot and children with autism spectrum disorder Tech. Rep. GAIPS-TR-001-19, GAIPS/INESC-ID, March 2019.

You may obtain any of the GAIPS technical reports from the corresponding author or on the group's web page (`http://gaips.inesc-id.pt/`).

gaips

INTELLIGENT AGENTS AND SYNTHETIC CHARACTERS GROUP