ELE - A Conversational Social Robot for Persons with Neuro-Developmental Disorders

Davide Fisicaro(✉), Franca Garzotto, Mirko Gelsomini, and Francesco Pozzi

Politecnico di Milano, P.zza Leonardo da Vinci, Milan, Italy
{davide.fisicaro, franca.garzotto, mirko.gelsomini, francesco.pozzi}@polimi.it

Abstract. Several studies explore the use of social robots in interventions for persons with cognitive disability. This paper describes ELE, a plush social robot with an elephant appearance that has been designed as a conversational companion for persons with Neuro-Developmental Disorders (NDD). ELE speaks through the live voice of a remote caregiver, enriching the communication through body movements. It is integrated with a tool for automatic gathering and analysis of interaction data that support therapists in monitoring the users during the experience with the robotic companion. The paper describes the design and technology of ELE and presents an empirical study that involved eleven persons with NDD using the robot at a local therapeutic center. We compared user engagement in two story-telling experiences, one with ELE and one with a face-to-face human speaker. According to our results, the participants were more engaged with ELE than with the human storyteller, which indicates, although tentatively, the engagement potential of conversational social robots for persons with NDD.

Keywords: NDD (Neurodevelopmental Disorder) · Disability · Social robot · Conversational companion · Engagement · Storytelling

1 Introduction

Social robots are characterized by the capability of communicating and interacting with users in a social and engaging manner [17, 18]. Several studies explore the use of social robots in interventions for persons with Neuro-Developmental Disorders (NDD). NDD is a general term for a group of conditions with onset in the developmental period [1] that are associated primarily with the functioning of the brain and the neurological system and are characterized by impairments in the personal, social, academic, and occupational spheres. Examples of NDD include Attention-Deficit Hyperactivity Disorder (ADHD), Autism Spectrum Disorder (ASD) and Down Syndrome.

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-29381-9_9) contains supplementary material, which is available to authorized users.

Published by Springer Nature Switzerland AG 2019
D. Lamas et al. (Eds.): INTERACT 2019, LNCS 11746, pp. 134–152, 2019.
https://doi.org/10.1007/978-3-030-29381-9_9

Social robots are thought [2, 7, 38–40] to elicit specific, desirable behaviors among persons with NDD, and to have the potential of promoting specific social skills. It is easier for persons with NDD to interact with robots than with humans because the former creates situations in which they can practice and learn in a safer and more pleasant manner [6, 9, 16]. Robots enable forms of embodied interactions that are appealing for these persons. They can offer human-like social cues (e.g., speaking, smiling) while maintaining object-like simplicity (e.g., limited facial expressions) and more generally provide sensory stimuli that are more predictable, less complex and less confusing [5, 7, 16].

The paper describes the design, implementation, and evaluation of ELE, a plush social robot that is intended to be used as a conversational companion during regular therapies for persons with NDD. ELE speaks through the digitally modified live voice of a remote caregiver and enriches verbal communication through body movements. The robot is integrated with a tool for automatic gathering and analysis of interaction data to support therapists' monitoring of the users during the experiences with ELE. We performed an empirical study at a local therapeutic center that aimed at evaluating the potential of ELE as a conversational companion for persons with NDD. The study involved eleven persons attending the center and their therapists. We compared participants' engagement in a story-telling experience with ELE against a similar one performed with a co-located human speaker. The results, although preliminary, show that participants were more engaged with ELE than with the face-to-face human storyteller, and indicate that conversational robots have a potential to increase engagement and motivation in interventions for subjects with NDD.

2 Related Work

In the last years, many researchers have investigated the application of social robots for persons with NDD, mainly considering children with autism (e.g., [3, 15, 16, 18–24, 30, 38–40]). Differently from other devices such as computers, tablets, and smartphones, social robots can engage children in the real world physically and emotionally, and offer unique opportunities of guided, personalized, and controlled social interaction and learning tasks [25]. Many social robots used in NDD interventions are remotely controlled by caregivers [16, 22, 26]. Autonomous behavior is implemented only in few cases, to support a specific learning task such as imitation, attention, communication, question-answering ability [27–29].

Several researchers explore the physical and dynamic characteristics of robots in relationship to subjects with NDD. Different shapes have been studies, e.g., "abstract", cartoon-like, or simplified or realistic human-like [3]; research suggests that individuals with NDD tend to prefer for something that is clearly "artificial" with respect to agents that have human-inspired characteristics [6, 31]. Some authors suggest that the shape of the robot should evoke a familiar element, such as a toy that the subject likes, or a cartoon character. For example, Teo [28] - a robot designed specifically for children with autism - resembles the popular cartoon characters of Minions. Puffy [7, 40] - an egg-shaped, inflatable, soft, and mobile robot – is inspired to Baymax, the inflatable healthcare robot of the popular movie Big Hero 6. The research reported in [32] and [33] shows that subjects with NDD may respond faster when cued by a robotic movement than human movement, and some social robots used in NDD therapy can move body parts [15, 22, 34, 35] or the entire body [7, 28, 40].

Several social robots exploit emotional features that seem to benefit children with NDD. Puffy [7] supports multisensory stimuli and multimodal interaction, can interpret child's intentions and emotion from her gestures and facial expressions, and can communicate emotions using sound and speech-based utterances, movements in space, and lights and projections embedded in its body. Keepon [22, 26] is a creature-like robot that is capable of expressing its attention (directing its gaze) and emotions (pleasure and excitement). An empirical study with autistic children showed that this robot triggered a sense of curiosity and security; the subjects spontaneously engaged in dyadic interaction with it, and then shared with the adult caregiver's the pleasure and the surprise they found in Keepon. KISMET [36] is an emotional robot which exploits eyebrows, ears, and mouth movements to expresses emotions depending on the way a human interacts with the robot. Teo [2, 41] supports the user's manifestation of emotions through the personalization of the robot body. Teo is equipped with a set of a detachable pieces like eyes, eyelids, or mouths that can be attached to its body and enables children to create "face" expressions. Its sensorized body can distinguish among caresses, hugs, and two levels of violent punches or slaps; the robot reacts to the different type of manipulations and expresses corresponding emotional states – happiness, angriness, or fear - using light, sound, vibrations, and movements.

3 The Design of ELE

ELE is a social robot that speaks through the live voice of a remote caregiver and enriching the communication using non-verbal signals, i.e., the movements of its body (i.e. trunk, eyes, ears). The goal of ELE is to play as a conversational companion, engaging users in dialogues or story-telling or story-listening tasks.

ELE is intended to address several needs: to mitigate a person's stress during therapies; to create moments of fun and trigger engagement; to enable the verbal communication between a person with NDD and a remote caregiver (e.g., therapist or educator) in situations when the former is unable to leave his/her home or to receive specific on-site intervention; to improve communication skills; to help therapists monitor the person's behavior remotely and automatically gather data that can be helpful for therapeutic and assessment purposes.

ELE is inspired by Huggable [10], a robotic teddy bear developed at MIT and used as a conversational mediator between hospitalized children and their caregivers. With Huggable, the remote operator can listen to the user via microphones; he/she can talk to the child by typing on a mobile device the text for the robot to speak, or by interacting directly with the robot through the embedded speaker (voice deformation features make the operator's voice not recognizable).

ELE has similar capabilities but provides some original features.

(1) It is integrated with a powerful application for therapists that enables caregivers to control the body behavior of the robot (not only eyes), i.e., movements of ears, eyes and trunk (and associated sounds). These therapist-controlled body expressions during a session of dialogue convey a personality-rich character to ELE. In addition, they are means to offer contextualized non-verbal backchanneling during the conversation. The term *backchannel* is used in linguistic to denote phatic expressions, primarily serving a social or meta-conversational purpose (e.g., to assess or acknowledge what is being said, to express attention or interest). The importance of verbal and non-verbal backchanneling in human-robot interaction has been demonstrated by a study using the robot Tega [12]. This study shows that humans are more engaged when interacting with a robot that is able to move (e.g., nodding) according with the semantic behind the conversation rather than a robot moving randomly.

(2) The application integrated with ELE also supports the transmission of both the audio and the *video* stream to and from the robot and enables to monitor the behavior of the person using ELE remotely. Therapists can monitor the user during the conversation, viewing the video recording and visualizing the evolution of the NDD person's emotions that are detected automatically from the streamed video. The same information can be visualized and inspected after a session, which is helpful for therapists to reflect on the person's behavior, inspect his or her progression, and tune future interventions.

(3) The design process of ELE has taken into account the replicability of the robot at *affordable* cost, which is an important issue for the adoption of a technology in real-life settings. Rather than building a new toy fully from scratch and at a high cost, as it happens for Huggable, we created ELE by reusing a cheap commercial plush toy provided with body movement capability and equipped it with off-the-shelf devices for input/output, control and connectivity devices. The result is a smart toy that can be developed, and replicated, at an affordable cost.

The first step in our design process was to identify a commercial plush toy that addressed the above requirements. Our choice was based on some needs pinpointed by NDD therapists we collaborate with:

- small size, because ELE is intended to be used also with children [3, 15];
- low cost;
- a shape different from the ones of typical pet toys (e.g., a dog, a cat, a teddy bear) to make the robot a unique entity;
- movement capabilities;
- colors that have positive psychological qualities [14] and are not problematic for persons with visual impairments.

We analyzed the rich catalogue offered by a worldwide known toy manufacturer - Giochi Preziosi - and finally selected a stuffed elephant (Fig. 1) sized $14 \times 7 \times 5.5$ in. Embedded in this commercial toy there are motors to move ears, trunk, and eyes. Movements of eyes and ears exploit a single DC (Direct Current) motor that rotates forward and backward. Another DC motor is used for moving the trunk up and down.

The skin colors are pale blue (most of the body), and, in some specific body parts, pale grey, pink and white. Blue is the color of the sky and sea; it is often associated with depth and stability and symbolizes trust, loyalty, wisdom, confidence, intelligence, and truth. Pink is thought to have a calming effect on people. Pale grey is considered to be a non-deflective neutral color that provokes neither a positive nor negative reaction [14].

Fig. 1. Front and left view of ELE.

We equipped the commercial toy with input/output devices and an embedded board. Input devices consists in a microphone and a camera, while the output device is a non-amplified speaker. The speaker and the microphone are positioned inside ELEs body, and the camera is positioned over ELE's head and hidden by a hat. These components as well as the native motors are controlled by an embedded board that also manages the communication with the dashboard application for the remote therapist.

4 Dashboard for Therapists

Therapists can control ELE remotely through an application called *dashboard* available on PC, tablet, or smartphone. The dashboard is accessible via a web page hosted on ELE's *internal* web server, automatically loaded when the system is turned on. The visual interface has been co-designed with the therapists participating in the project and its usability was evaluated during an empirical test with external therapists from a different therapeutic center.

A control panel (Fig. 2-right side) enables therapists to control ELE movements and define the audio settings. Predefined combination of sounds and movements, defined with therapists, are provided to facilitate control and provide backchanneling during the conversation. An example is the "Trumpeting": while ELE reproduces a bellow, it moves its trunk up-down-up and ears move back and forth one time. Audio/video stream can be activated by the operator at any time.

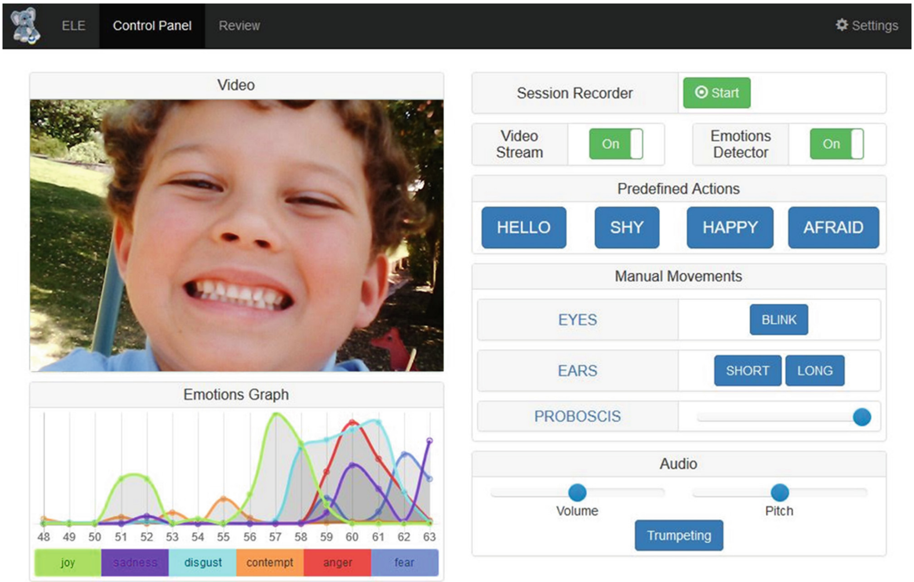

Fig. 2. Dashboard as seen by the remote operator during a session. On the left side, video stream and emotion graph. On the right side, ELE's control (voice modification and motors).

The progression of emotions extracted from the video are visualized as graphs on a separate visualization panel, which also offers a "filter" function to select a subset of the available emotions to display (Fig. 3).

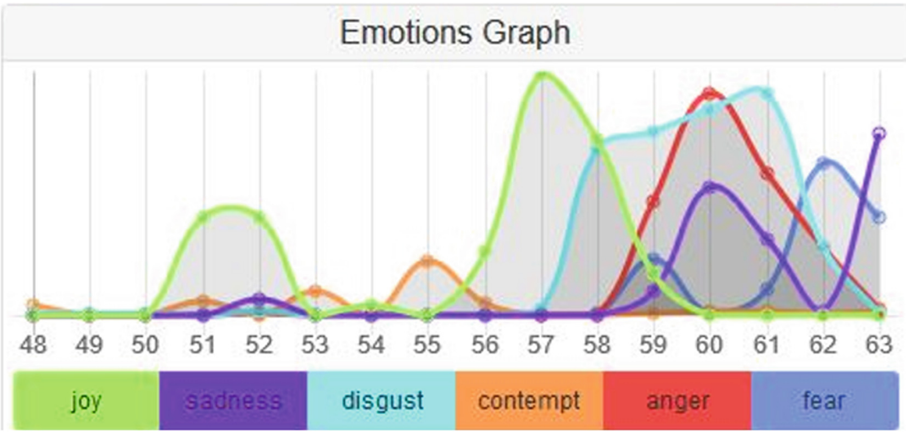

Fig. 3. Focus on the emotions graph.

At the end of the session, therapists can save all emotional information as well as with the audio/video streams, and the movements of ELE performed during the interaction. All data are locally stored in the therapist's device and are accessible at any time (Fig. 4).

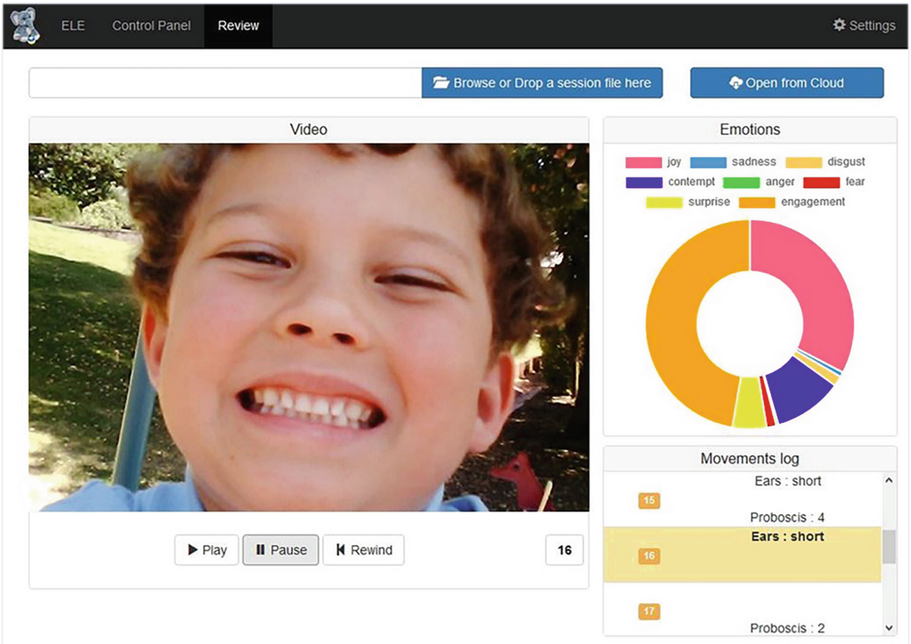

Fig. 4. Review section of the dashboard. On the left side, the recorded video. On the right side, summary of recorded emotions and log of ELE's action.

5 Technology

5.1 Hardware Features

All input/output components and motors are connected to a Raspberry Pi 3 Model B board (RaspberryPi Foundation - https://www.raspberrypi.org) running Raspbian Stretch Lite operating system. The power consumption is very low, and a commercial power bank is placed inside ELE to supply the whole system, giving almost five hours of autonomy. A retractable USB cable is placed on the back in order to charge the power bank.

5.2 Software Architecture

The web application implementing the dashboard for therapists exploits a web server that is hosted inside ELE itself and is automatically loaded when the system is turned on. Network connection between ELE and the application is handled by a proxy.

The software module that manages the connection, as well as the modules used to manage the commands received from the operator activating ELE's motors and the other input/output devices, are coded in Python. Sensible data (as well as all the other data) from ELE to the therapist's device are transferred over HTTPS protocol and all data are encrypted. A full-duplex audio/video streaming is opened between the operator and the NDD person using ELE; a specific component is devoted to modifying the therapist's voice.

The real-time emotion analysis component –coded in JavaScript - exploits the ELE user's person's facial expression retrieved from the camera streaming is performed on the therapist's browser. Facial expression analysis relies on *Affdex*, a commercially available recognition software [11]. The software collects information about facial position and facial key points to provide a value between 0 and 100 every 100 ms. This value represents the probability that the subject is feeling a particular emotion; emotions set includes joy, sadness, disgust, contempt, anger and fear. Every second, an average of ten values for each emotion is calculated and the generated sequences are displayed in real-time via the emotion graph in the control panel. This information is logged, saved and synchronized with the video recording of the session and the movements ELE did.

6 Empirical Study

6.1 Goal

We performed an exploratory empirical study at a local care center to investigate the potential of ELE as a conversational companion for persons with NDD. We focused on a specific aspect, namely, the robot capability of promoting *engagement* among this target group. To this end, we compared the engagement potential of ELE against the one of a therapist speaking face-to-face.

Engagement is a broad concept and there is limited agreement on the definition and operationalization of the construct. We embraced the definition from Chapman ([4], p. 3), who defines engagement as *"…something that draws us in, that attracts and holds our attention."*. Engagement is widely acknowledged learning *facilitator* [12] and, for subjects with NDD, it has an even stronger role. The impairments associated to NDD create a persistent state of insecurity and uncertainty, a tendency to withdrawal and self-inhibition, and a difficulty to stay focused on something for a prolonged time, which hinders the willingness and capability to be involved in a task and to act upon the associated objects. Among subjects with NDD, reaching and maintaining a state of engagement is a *precondition* for any learning process to take place [37].

6.2 Engagement Metrics

Prior work with persons with cognitive impairments [42] suggests that *gaze* is a good quantitative indicator of engagement. For the purpose of our study, we assumed that the person with NDD – hereinafter referred to as "*participant*" - is *engaged* when *he/she* *looks at* the face of the "*speaking agent*", i.e., the on-site therapist (in experimental situation S1) or ELE (in experimental situation S2).

Engagement metrics were defined using the terms and expressions reported below.

- *Area of interest:* the *face* of the speaking agent (ELE or the face-to-face therapist);
- *Total Session Time (T_{tot}):* total duration of an experimental session;
- *Focus Interval:* an interval of time during a session in which the participant *maintains* her gaze on a point within the area of interest;
- *Total Focus Time (T_f):* the sum of all Focus Intervals during a session, i.e., the total time during which the participant looks at the area of interest during a session.
- *Total Focus Change (C):* the number of times the participant *moves* her gaze from *outside* to *inside* the area of interest.

Areas of interest are calculated starting from simplified geometric models of ELE's and the therapist's face. ELE's face is modeled as a circumference while the human speaker's one as an ellipsis where the two principal axis are the head width and height.

We used a commercial eye tracker (https://steelseries.com) to gather gaze measures. The data retrieved from the eye tracker are sequences of tuples containing information on where the subject is looking at (x and y coordinates referred to a reference system whose origin is on the setting) at a specific time instant. These data are aggregated in order to measure Focus Intervals: a tuple belongs to a Focus Interval when its coordinates are inside the area of interest.

T_f and C are calculated from the set of Focus Intervals for each participant and each session, and are then normalized with respect to the session duration to obtain the following two measures:

- *Performance on Focus Change:* $P_{fc} = \frac{C}{T_{tot}}$
- *Performance on Focus Time:* $P_f = \frac{T_f}{T_{tot}}$

The *Performance on Focus Time* can be interpreted as the *probability* that at a given time instant the subject is looking at the area of interest.

We can assume that an *increase* of *Performance on Focus Time* indicates an *increase of engagement*. Still, this single measure should be considered together with *Performance of Focus Change*: for instance, if a subject has the same P_f in two sessions but P_{fc} increases in the second session, the latter session should be considered *less* engaging because the duration of the single intervals of focus on the area of interest decreases. For this reason, we assume that a *decrease* in *Performance on Focus Change* indicates an *increase of engagement*.

This information is used to create a scoring function $E(P_f, P_{fc})$ called *session performance* that give us the possibility to compare two sessions for a single participant. This value also represents the engagement.

Assuming that $E = f(P_f) \circ g(P_{fc})$, where f and g are two unknown functions and $\circ$ an unknown operator between the two functions, we must define f, g and $\circ$ in order to perform the scoring. Since we want to maintain proportionality, we consider the operator "$\circ$" as a multiplication: at this point $E = f(P_f)g(P_{fc})$.

Some considerations on the nature of functions f and g should be made: both of them must be defined, at least, in the interval [0, 1], because data are positive and can be null. Furthermore f must be an increasing monotone function because its argument, P_f, is directly proportional to E, while g must be a decreasing monotone function because its argument, P_{fc}, is inversely proportional to E. Having in mind the above constraints, the choice fell on a negative exponential function for g, while f is linear. The scoring function E is defined as follows:

$$E\left(P_f, P_{fc}\right) = P_f * e^{-P_{fc}}$$

This definition must be completed to take into account some particular cases:

- when $P_f = 0$ (the participant never looked in the area of interest), $E = 0$;
- when $P_{fc} = 0$ there are two possibilities:
  - the participant never looked inside the area of interest, then $E = 0$;
  - the participant always looked inside the area of interest without exiting from it, then $E = P_f\, e^0 = P_f$. Since $P_f = 1$ when the participant always looks in the area of interest, then $E = 1$;
  - when $P_f = 1$ (the maximum value), then $P_{fc} = 0$ and $E = 1$ (see previous point);
  - when P_{fc} grows, tending to infinite, the number of interactions is very high but the permanence inside the area of interest for each interaction is almost 0, then E should be equal to 0. In fact, $\lim_{P_{fc} \to \infty} P_f * e^{-P_{fc}} = 0$ due to the negative exponential;
  - for any other value of P_{fc} (which is always positive, without considering the already discussed case in which $P_{fc} = 0$), $e^{-P_{fc}} < 1$ and $0 < P_f < 1$ by definition, so $0 < E < 1$.

In summary, the completed definition of the scoring function is the following:

$$E\left(P_f, P_{fc}\right) = P_f * e^{-P_{fc}}, \text{ with } 0 \leq E \leq 1.$$

6.3 Participants

Finding a homogeneous group of participants and controlling bias introduced by individual differences is acknowledged as very difficult in any study involving persons with NDD. We involved 11 participants recruited among the persons attending the center where the study took place. Their age range was 25–43 ($\mu = 31.09$, $\sigma = 5.1$). The group was heterogeneous with respect to their diagnosis but homogenous with respect to intellectual functioning level (medium), as described in Table 1.

Table 1. Subject's age, gender (G) and diagnosis

Subject	Age (years)	G	Diagnosis
1	35	F	Mental retardation of medium degree with severe limitations of personal autonomy, polyvoltine syndrome, obesity
2	28	F	Mental retardation of medium-severe entity associated with behavioral disorders symptomatic of cerebral malformation
3	30	M	Mental retardation in genetic syndrome with minor malformative aspects
4	25	M	Autism spectrum disorder
5	27	F	Mental retardation of medium degree, global hypoututism, growth hormone deficiency, diabetes mellitus
6	28	M	Spastic dysplasia with a medium to severe cognitive delay as outcomes of neonatal distress
7	36	F	Severe mental insufficiency and deficit psychosis
8	31	M	Cornelia de Lange syndrome with serious mental retardation
9	29	M	Mental retardation of medium degree with behavioral disorders
10	30	F	Borderline personality disorder with behavioral abnormalities in mental retardation with very limited relational skills
11	43	F	Average oligophrenia on a cerebropathic basis

6.4 Method

The study respected the ethical rules and procedures required by our university and the study was approved by our Ethical Committee and the one at the care center where we performed the study took. The head of the therapeutic team at the center identified a set of potential participants, explained them the study, and asked them if they wanted to participate. The same procedure was carried out among therapists and families. An informed consent (also including data treatment rules) was signed by parents/guardians. All digital data were anonymized and were stored in a certified secure server, while paper documentation was kept in a dedicated lock-room.

The study had a two-conditions within-subjects design. Each participant attended *two sessions*, and experienced a different experimental condition in each session: S_1 = *"Talking with a face-to-face human speaker"* and S_2 = *"Talking with ELE"*. The order of conditions was *randomized* among participants.

Instruments, stimulus, set-up, and test protocol (time of exposition to the stimulus, physical distance between the participant and the speaking agent, room setting, session scheduling) were defined carefully to control as much as possible for the many potentially confounding variables. The most challenging requirement was to standardize the stimulus in the two experimental conditions (human speaker and ELE), minimizing the differences in content and voice features.

Since storytelling is a frequent activity in interventions among persons with NDD [8], we created two short tales that had the same number of words, similar plot and environment, same number of characters, resolution pattern and duration. According to therapists, they could be regarded as "equivalently engaging". They were randomly assigned to each experimental condition. The *same therapist* (unknown to all participants) told the story through ELE and face-to-face. To exclude differences due to individual voice characteristics, ELE spoke using the *pre-recorded reading of the story;* the therapist's voice was digitally modified to hide its human nature and simulate the voice of a fantasy character. The therapist was trained to tell the other story trying to use the same voice tone and speech rhythm as much as possible.

The experimental sessions took places in a dedicated room. The setting included two tables, some chairs, very neutral furniture (bookshelf, baskets) and a frame, and remained the same for all sessions. The frame had the purpose of hiding the technological instruments (an eye tracker to collect the fixation point and a camera to record the session). The frame represents a simple natural landscape, which was designed with the help of a therapist and was used in all sessions. (see Fig. 5). The frame was placed on the table, with the participant sitting in front of it, while the speaker was placed behind the subject (Figs. 6 and 7). The chair was located between 50 and 80 cm from the frame (otherwise his/her gaze would not be detected by the eye tracker). a psychologist and a member of the design team participated as observers, taking notes during sessions.

Each session followed the same protocol:

1. Before entering the study room, the participant is informed on what is going to happen ("You will enter, sit on the chair placed in front of the table where a nice frame is placed, and listen to a story") and the two observers are introduced to him/her.
2. The participant enters the room and is guided to the chair. The two observers move behind the participant, out of the camera vision angle.
3. The participant listens to the story; when it ends, he/she is invited to cheer the speaker and is accompanied out of the room.

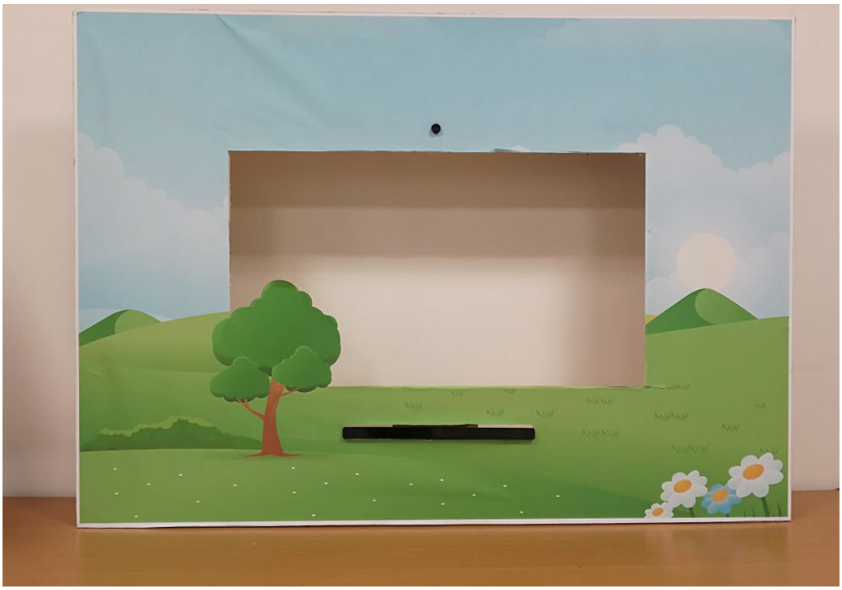

Fig. 5. Frame used in the study.

Fig. 6. Session with human speaker (S1).

Fig. 7. Session with ELE as a speaker (S2)

6.5 Main Results

The analysis of the data collected during the study does not consider participant #11 (aged 43) who attended one session only. In the rest of this section, we will use the expression "*sessions i*" (i = 1, 2) to indicate sessions with experimental condition S_i.

The results concerning *Performance on Focus Time* (P_f) and *Performance on Focus Change* (P_{fc}) in the two types of sessions –are summarized in Tables 2 and 3.

Table 2. Data gathered during Sessions 1 - with human speaker

Subject	P_f [%]	P_{fc} (ms^{-1}) [%]
1	6.73	0.05
2	50	0.5
3	0	0
4	0.1	0.01
5	6	0.01
6	16	0.70
7	70	0.03
8	0	0
9	37	0.02
10	0	0

Table 3. Data gathered during Sessions 2 - with ELE

Subject	P_f [%]	P_{fc} (ms^{-1}) [%]
1	66	0.1
2	31	0.3
3	6	0.1
4	95	0.01
5	49	0.05
6	10	0.05
7	97	0.02
8	49	0.05
9	83	0.08
10	98	0.02

Participants' Scores in the two experimental conditions, and respective variations, are shown in Table 4 and visualized in Figs. 8 and 9.

Table 4. Scores and score differences in the two experimental conditions

Subject	Scoring (E) [%]		$\Delta E = E_2 - E_1$ [%]
	Session 1 (E_1)	Session 2 (E_2)	
1	6	66	60
2	50	31	−19
3	0	5.8	5.8
4	0.1	95	94.9
5	6	48	42
6	17	10	−7
7	70	97	27
8	0	49	49
9	37	83	46
10	0	98	98

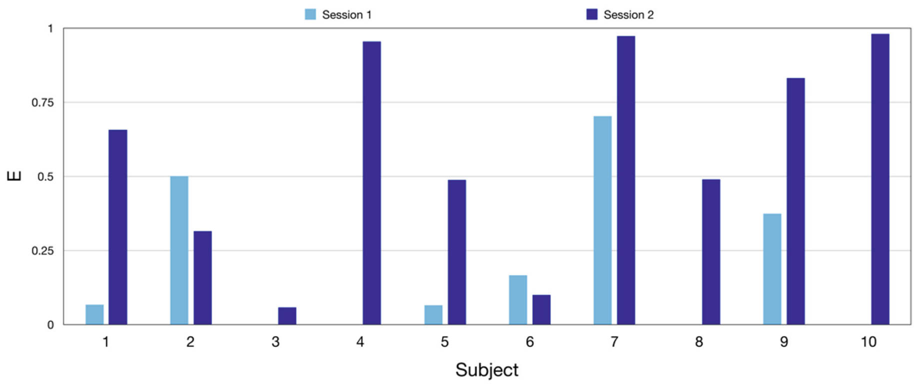

Fig. 8. Absolute scores for each participant in each experimental condition

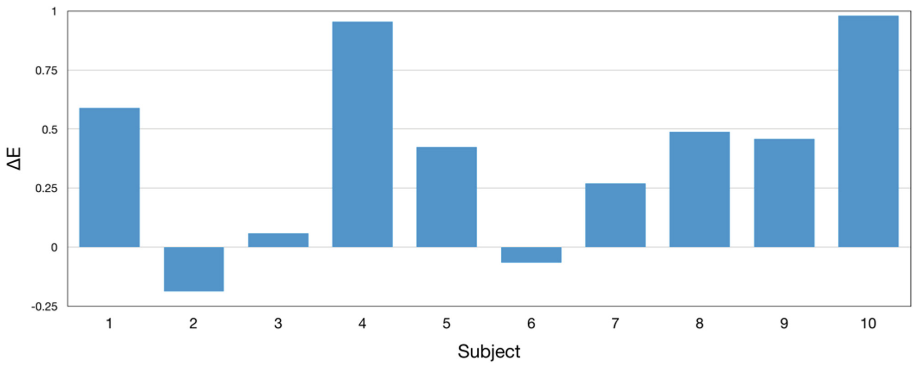

Fig. 9. Variations of the scores for each participant in each experimental condition

6.6 Discussion

The main finding is that for 8 participants out of 10 *(80%)* we measured a higher engagement score in the session with ELE than in the ones with the face-to-face therapist. This result may encourage the future adoption of ELE as conversational companions in interventions for persons with NDD at the therapeutic center where the study was performed. Still, our study is exploratory and has a number of limitations.

We involved the participants in one session for each type of stimulus (ELE and face to face human speaker). A longer exposure to both experimental conditions would increase the validity of our results. We may wonder if the positive results for ELE could be ascribed to the "novelty effect" of the robot. The answer is "probably not". For all subjects involved in our study (like for many people with NDD), the "unknown" is often a source of distress and discomfort and these persons tend to manifest rigidity towards any new situation. Novelty therefore should not be considered a facilitator of engagement.

The number of participants (11) was relatively small, although this sample size is similar to many studies on robots for persons with NDD [2, 16, 23, 30]. The main critical issue is the variability between the participants: even if their cognitive level is comparable to neuro-typical children aged 8–10, they have relevant differences in of age, impairments, severity of the disorders, capability of processing and reacting to sensory stimuli.

Participant variability is a typical problem in empirical research among persons with disability and are one of the main reasons why statistical methods can hardly be applied in this field. Particularly, the descriptive analysis of our data indicates that the distance from the Gaussian Curve was too high to allow an inferential analysis, and we opted to present individual data and results by participants.

The individual variability also makes difficult to generalize the results of our study: persons with NDD different from our participants, or the same persons experiencing ELE and the human speaker in different ambient conditions, may not manifest similar engagement/disengagement trends as in our study.

7 Conclusion

We have presented the design, technology, and evaluation of ELE, a novel plush social robot with an elephant appearance that has been designed as a conversational companion for persons with NDD. ELE speaks through the live voice of a remote caregiver and is integrated with a tool for automatic gathering and analysis of interaction data.

Even if ELE is inspired to an existing smart toy that offers conversational proxy facility (MIT Huggable [43]), our research has some peculiar features that makes it an interesting contribution to research on social robots for persons with NDD.

In ELE, the physical design per se is not original as we reused a commercial toy. Still, “smartifying” an existing toy rather than building a new one from scratch (like Huggable) has an advantage in terms of development cost and therefore in terms of the potential for adoption.

The dashboard for caregivers integrated with ELE offers more features than Huggable’s telecontrol application, and enables monitoring, analyzing and visualizing various kinds of user’s data including emotion flows.

An additional contribution of our research is the empirical exploration of the engagement potential of conversational social robots. The operationalization of engagement and its metrics are novel. To our knowledge, conversational social robots were never studied among persons with cognitive impairments. Huggable for example was designed for and tested with neurotypical hospitalized children and was evaluated in terms of effects on user relaxation and communication. Engagement of smart toys for persons with NDD was explored in one previous work only [44] that used non-conversational e-toys and weaker metrics for engagement evaluation. Our initial empirical study suggests that ELE might be a more engaging conversational companion for persons with NDD than human speakers. This outcome needs to be validated in future research. Engagement is a necessary precondition for any learning process to take place among our target group. If our results are confirmed, ELE could be used as a complement to traditional interventions for this target group, e.g., to promote verbal communication skills. In addition, since reaching a state of engagement is known to help releasing tension, ELE could be used to alleviate a person’s stress during any therapy. An additional benefit of ELE that will deserve future research is related to the use of the robot for remote therapy, where a distant caregiver provides verbal interventions remotely. In this respect, ELE can be a useful tool for persons with NDD who cannot leave home.

A further direction for future research concerns the data collected by ELE. Initial feedbacks by therapists pinpoint the utility of the visualization and analysis tools for monitoring the emotional expression and communication attitude of the persons with NDD who use ELE. Still, little is known about the way therapists could use this information to improve their interventions. Finally, the audio and video streams collected by ELE represent a wealth of information on the behavior of persons with NDD: they could be analyzed with appropriate art AI tools and exploited for NDD diagnosis.

Acknowledgements. The authors are grateful to Gianmarco Giummarra, Salvatore Ferrigno and Mattia Melli for their contribution to the development of ELE and to the execution of the empirical study. We warmly thank all persons at the therapeutic center Fraternità e Amicizia (Milan) who participated in our research.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub., Arlington (2013)
2. Bonarini, A., Garzotto, F., Gelsomini, M., Romero, M., Clasadonte, F., Yilmaz, A.N.C.: A huggable, mobile robot for developmental disorder interventions in a multi-modal interaction space. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 823–830. IEEE (2016)
3. Cabibihan, J.J., Javed, H., Ang, M., Aljunied, S.M.: Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. Int. J. Social Robot. **5**(4), 593–618 (2013)
4. Chapman, P.M.: Models of engagement: intrinsically motivated interaction with multimedia learning software. Ph.D. thesis, University of Waterloo (1997)
5. Dautenhahn, K., Werry, I.: Issues of robot-human interaction dynamics in the rehabilitation of children with autism. Proc. Anim. Animats **6**, 519–528 (2000)
6. Dautenhahn, K., Werry, I.: Towards interactive robots in autism therapy: background, motivation and challenges. Pragmat. Cogn. **12**(1), 1–35 (2004)
7. Garzotto, F., Gelsomini, M., Kinoe, Y.: Puffy: a mobile inflatable interactive companion for children with neurodevelopmental disorder. In: Bernhaupt, R., Dalvi, G., Joshi, A., Balkrishan, D.K., O'Neill, J., Winckler, M. (eds.) INTERACT 2017. LNCS, vol. 10514, pp. 467–492. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67684-5_29
8. Garzotto, F., Paolini, P., Sabiescu, A.: Interactive storytelling for children. In: Proceedings of the 9th International Conference on Interaction Design and Children, IDC 2010, pp. 356–359. ACM, New York (2010)
9. Huijnen, C.A., Lexis, M.A., Jansens, R., de Witte, L.P.: Mapping robots to therapy and educational objectives for children with autism spectrum disorder. J. Autism Dev. Disord. **46** (6), 2100–2114 (2016)
10. Jeong, S., et al.: Designing a socially assistive robot for pediatric care. In: Proceedings of the 14th International Conference on Interaction Design and Children, pp. 387–390. ACM (2015)
11. McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., Kaliouby, R.e.: AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA 2016, pp. 3723–3726. ACM, New York (2016)
12. O'Brien, H.L., Toms, E.G.: What is user engagement? A conceptual framework for defining user engagement with technology. J. Am. Soc. Inform. Sci. Technol. **59**(6), 938–955 (2008)
13. Park, H.W., Gelsomini, M., Lee, J.J., Breazeal, C.: Telling stories to robots: the effect of backchanneling on a child's storytelling. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 100–108. ACM (2017)
14. Pauli, D., Egerton, J., Carpenter, B.: The sunfield colour impact project. PMl. Dl. inl **12**(2), 12–16 (1999)
15. Robins, B., Dautenhahn, K., Te Boekhorst, R., Billard, A.: Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Univ. Access Inf. Soc. **4**(2), 105–120 (2005)
16. Scassellati, B., Admoni, H., Matarić, M.: Robots for use in autism research. Annu. Rev. Biomed. Eng. **14**, 275–294 (2012)
17. Feil-Seifer, D., Matarić, M.J.: Socially assistive robotics. IEEE Robot. Autom. Mag. **18**(1), 24–31 (2011)

18. Stiehl, W.D., Lieberman, J., Breazeal, C., Basel, L., Lalla, L., Wolf, M.: Design of a therapeutic robotic companion for relational, affective touch. In: IEEE International Workshop on Robot and Human Interactive Communication, ROMAN 2005, pp. 408–415. IEEE, August 2005
19. Chella, A., Barone, R.E., Pilato, G., Sorbello, R.: An emotional storyteller robot. In: AAAI Spring Symposium: Emotion, Personality, and Social Behavior, pp. 17–22, April 2008
20. Den Brok, W.L.J.E., Sterkenburg, P.S.: Self-controlled technologies to support skill attainment in persons with an autism spectrum disorder and/or an intellectual disability: a systematic literature review. Disabil. Rehabil. Assistive Technol. **10**(1), 1–10 (2015)
21. Diehl, J.J., Schmitt, L.M., Villano, M., Crowell, C.R.: The clinical use of robots for individuals with autism spectrum disorders: a critical review. Res. Autism Spectr. Disord. **6** (1), 249–262 (2012)
22. Kozima, H., Michalowski, M.P., Nakagawa, C.: Keepon. Int. J. Social Robot. **1**(1), 3–18 (2009)
23. Michaud, F., Duquette, A., Nadeau, I.: Characteristics of mobile robotic toys for children with pervasive developmental disorders. In: 2003 IEEE International Conference on Systems, Man and Cybernetics, vol. 3, pp. 2938–2943. IEEE, October 2003
24. Ricks, D.J., Colton, M.B.: Trends and considerations in robot-assisted autism therapy. In: 2010 IEEE International Conference on Robotics and Automation (ICRA), pp. 4354–4359. IEEE, May 2010
25. Short, E., et al.: How to train your DragonBot: socially assistive robots for teaching children about nutrition through play. In: 2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, pp. 924–929. IEEE, August 2014
26. Kozima, H., Nakagawa, C., Yasuda, Y.: Children–robot interaction: a pilot study in autism therapy. Prog. Brain Res. **164**, 385–400 (2007)
27. Robins, B., et al.: Scenarios of robot-assisted play for children with cognitive and physical disabilities. Interact. Stud. **13**(2), 189–234 (2012)
28. Bonarini, A., Clasadonte, F., Garzotto, F., Gelsomini, M.: Blending robots and full-body interaction with large screens for children with intellectual disability. In: Proceedings of the 14th International Conference on Interaction Design and Children, pp. 351–354. ACM, June 2015
29. Kory Westlund, J., et al.: Tega: a social robot. In: The Eleventh ACM/IEEE International Conference on Human Robot Interaction, p. 561. IEEE Press, March 2016
30. Dautenhahn, K., Werry, I., Salter, T., Boekhorst, R.T.: Towards adaptive autonomous robots in autism therapy: varieties of interactions. In: Proceedings of the 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, vol. 2, pp. 577–582. IEEE, July 2003
31. Robins, B., Dautenhahn, K., Dubowski, J.: Does appearance matter in the interaction of children with autism with a humanoid robot? Interact. Stud. **7**(3), 509–542 (2006)
32. Bird, G., Leighton, J., Press, C., Heyes, C.: Intact automatic imitation of human and robot actions in autism spectrum disorders. Proc. Roy. Soc. Lond. B Biol. Sci. **274**(1628), 3027–3031 (2007)
33. Pierno, A.C., Mari, M., Lusher, D., Castiello, U.: Robotic movement elicits visuomotor priming in children with autism. Neuropsychologia **46**(2), 448–454 (2008)
34. Shamsuddin, S., et al.: Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO. In: 2012 IEEE 8th International Colloquium on Signal Processing and its Applications (CSPA), pp. 188–193. IEEE, March 2012
35. Shibata, T., Mitsui, T., Wada, K., Tanie, K.: Subjective evaluation of seal robot: paro-tabulation and analysis of questionnaire results. J. Robot. Mechatron. **14**(1), 13–19 (2002)

36. Breazeal, C.: A motivational system for regulating human-robot interaction. In: AAAI/IAAI, pp. 54–61, July 1998
37. Ferrara, C., Hill, S.D.: The responsiveness of autistic children to the predictability of social and nonsocial toys. J. Autism Dev. Disord. **10**(1), 51–57 (1980)
38. Colombo, S., Garzotto, F., Gelsomini, M., Melli, M., Clasadonte, F.: Dolphin Sam: a smart pet for children with intellectual disability. In: Proceedings of the International Working Conference on Advanced Visual Interfaces, pp. 352–353. ACM, June 2016
39. Tam, V., Gelsomini, M., Garzotto, F.: Polipo: a tangible toy for children with neurodevelopmental disorders. In: Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction, pp. 11–20. ACM, March 2017
40. Bonarini, A., Clasadonte, F., Garzotto, F., Gelsomini, M., Romero, M.: Playful interaction with Teo, a mobile robot for children with neurodevelopmental disorders. In: Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, pp. 223–231. ACM, December 2016
41. Hu, F. (ed.): Virtual Reality Enhanced Robotic Systems for Disability Rehabilitation. IGI Global, Hershey (2016)
42. Armstrong, T., Olatunji, B.O.: Eye tracking of attention in the affective disorders: a meta-analytic review and synthesis. Clin. Psychol. Rev. **32**(8), 704–723 (2012)
43. Stiehl, W.D., et al.: The huggable: a therapeutic robotic companion for relational, affective touch. In: ACM SIGGRAPH 2006 Emerging Technologies, p. 15. ACM, July 2006
44. Beccaluva, E.A., et al.: Exploring engagement with robots among persons with neurodevelopmental disorders. In: 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 903–909. IEEE, August 2017