

An Experimental Study of Embodied Interaction and Human Perception of Social Presence for Interactive Robots in Public Settings

Damith Herath, *Member, IEEE*, Elizabeth Jochum, and Evgenios Vlachos

Abstract— The human perception of cognitive robots as social depends on many factors, including those that do not necessarily pertain to a robot's cognitive functioning. Experience Design offers a useful framework for evaluating when participants interact with robots as products or tools and when they regard them as social actors. This study describes a between-participants experiment conducted at a science museum, where visitors were invited to play a game of noughts and crosses with a Baxter robot. The goal is to foster meaningful interactions that promote engagement between the human and robot in a museum context. Using an Experience Design framework, we tested the robot in three different conditions to better understand which factors contribute to the perception of robots as social. The experiment also outlines best practices for conducting human-robot interaction research in museum exhibitions. Results from the study indicate that perceived social presence can be evaluated using a combination of HRI and Experience Design methods that measure co-presence and co-experience.

Index Terms—Human Robot Interaction, HRI, Cognitive robotics, Social Computing, Experience Design

I. INTRODUCTION

SOCIAL robotics represents the natural evolution of robots, embodied and cognizant, centrally situated in social contexts with perceived social abilities. As such, any technical advancement needs to be informed by Human Robot Interaction (HRI) and ethnographic studies that are conducted in contextually relevant settings. Thus far, many studies are confined to laboratory conditions that lack the necessary 'social' aspects of authentic social interactions. This is understandable given the challenges of creating 'social' robots that are truly autonomous. As shown in [1, 2], the perceived success of an interaction with a robot could well depend on the environment and situational context. To fully understand how humans perceive the social presence of interactive robots, more experiments should be conducted in settings that afford natural and fluid interactions, while still allowing for thorough

observation and analysis. Benchmarks for studying HRI in dynamic contexts should account for overall perceived social presence, which may depend on factors beyond the functional capabilities of the robot.

Real-life interactions with cognitive robots are still rare occurrences for most people, and museums and science centers offer opportunities for the general public to interact with robots and learn about these emerging technologies. Museums are important sites for bringing humans and robots into closer contact, and have proved tractable sites for studying HRI in semi-controlled environments [3-8]. Museums allow researchers to study how humans naturally engage with interactive robots and perceive their motions and behavior, while still allowing for a controlled environment to carry out experiments in the wild. However, not much of this research has been applied to the design of permanent or semi-permanent robot exhibits: despite significant advancements in robotics, robots in museums are situated behind glass or in cages, and are physically separated from the public. Such "off-limits" exhibits further the perception of robots as functional devices to be avoided, not as social actors meant for interaction. Conventional approaches to designing robot exhibitions typically do not allow for meaningful interactions; however, museum curators can use exhibits to study HRI in greater detail and apply their findings to the development of diverse, interactive exhibitions.

Working together with museum curators, we designed an interaction scenario informed by theories from Experience Design (ED) for a permanent museum exhibit, and conducted several trials using HRI and ED methods. The results from this experiment will inform the design of a fully autonomous, interactive robot exhibit. We conducted ethnographic research and experimental studies to test which factors influence the perception of robots as social. The goal is to leverage the environmental and situational context afforded by museums to design engaging exhibitions that facilitate *co-experiences* between humans and robots. The broader scope of this study is to demonstrate how methods from ED might contribute to the study of human perception through the concepts of co-presence and co-experience.

Our experiment is unique in that we have incorporated ED methods to develop strategies that support HRI not only from a psychological standpoint, but also from the user experience perspective. ED explores aspects of playfulness, surprise, and

D. C. Herath is with the Human Centered Technology Research Center, University of Canberra, University Dr, Bruce ACT 2617 Australia (e-mail: damith.herath@canberra.edu.au).

E. Jochum is with the Department of Communication and Psychology, Aalborg University, Denmark (e-mail: jochum@hum.aau.dk).

E. Vlachos is with the Syddansk Universitetsbibliotek, University of Southern Denmark, Campusvej 55, 5230 Odense M, (e-mail: evl@bib.sdu.dk).

enchantment [9], and ED methods have been successfully incorporated into human-computer interaction research (HCI), which emphasizes human experience, perceived co-presence and the overall aesthetics of interaction. ED describes user-product interactions (how people interact with tools, or products) and dimensions of user experience (how interactions unfold, and emotion and experience is evoked) [10]. In HCI, the focus is primarily on user experience with interactive screen-based interfaces. The application of ED methods for evaluating perceived social presence for embodied robots is a novel contribution of this study. User-product (or in our case, human-robot) interactions can be characterized as *fluent* (automatic and well-learned), *cognitive* (requiring cognitive effort on the part of the human), and *expressive* (helping the user to form a relationship to a technological tool). Dimensions of user experience are defined as experience, an-experience, and co-experience. Examples of *Experience* are walking in a park, or using an instant messaging system, where users engage in “self-talk” while interacting with products. An *experience* has a beginning and an end, such as riding a rollercoaster, or watching a movie. *Co-experiences* create meaning and emotion for the user, usually by interacting with others, such as playing a game in a shared physical, or online space. These categories are relevant for HRI research because they provide a framework for evaluating when participants interact with robots as products, or tools and when they regard them as social actors. *Co-experience* is closely related to the perceived social presence of artificial agents.

In our observational study of current robot exhibits in museums, we observed that *co-experiences* usually take place between museum-goers and sometimes between visitors and museum guides, but rarely between visitors and robots. We observed that people were more likely to stay longer and engage in extended interactions with exhibits when there was another human present, such as a museum guide. For example, at a recent robot exhibition at the Science Museum in London, visitors gathered around a guide holding a robotic cat who described the robot's features and interacted with it, and invited the public to join her. This interaction contrasted with the more technologically sophisticated robots on display behind barriers, similar to passive exhibits found in natural history museums (Fig. 1). In these instances, co-experience is facilitated by the guide who acts as a mediator, contributing to the perception of the robot's social presence.

We devised an experiment to test different interaction scenarios with and without a human mediator's presence. Building on our prior work with the Powerhouse Museum (Sydney, AUS), we found that interactive games can facilitate meaningful interaction experiences. We decided to use the game of noughts and crosses that has been extensively used by interactive machines in museums [10]. Furthermore, [11] found that participants show a greater level of social



Fig. 1. Robot exhibition at the Science Museum (London), visitors observe robots from behind a barricade. The robots were interactive, however we observed that *co-experiences* typically occurred between museum goers, or between museum-goers and guides, but rarely between museum-goers and robots.

engagement and make greater attributions of mental states when playing against a robot that cheats, which prompted us to devise a cheat function whereby the robot disrupts the game flow play by placing a tile on one of the player's already placed tiles to prevent a win. In our experiment the participant engaged with the robot in one of three conditions where, (1) the game mediates the co-experience, (2) the researcher mediates the co-experience, and (3) the researcher mediates the co-experience, but the robot is cheating.

The site for our experiment was the National Science and Technology Centre - Questacon in Canberra, the Australian Capital Territory. Over a three-day period, the collaborative robot Baxter [12] was installed in one of the museum galleries and was programmed to play a game of noughts and crosses (tic-tac-toe) autonomously with a human participant. The cheat-function was implemented using Wizard of Oz. Using self-reporting and video analysis, we evaluated the participants' experience and perception of co-presence, combining questions from Lombard and Ditton [13] (also used in [15]) and the Nowak and Biocca questionnaire [14] for studying physical presence and social presence for virtual agents. The experiments revealed the potential of combining ED methods with HRI studies, and also the advantages and challenges associated with conducting human subject research in museum contexts.

II. RELATED WORK AND MOTIVATION

Humans have long been fascinated by the possibility of socially interactive machines [15]. With advances in various related technologies, it is now possible to create machines that adequately mimic social behaviors. Breazeal identifies a sub class of machines that people anthropomorphize in order to interact with them as social robots [16]. According to Fong et. al. [17], socially interactive robots tend to exhibit the following “human social” characteristics: express and/or perceive emotions, communicate with high-level dialogue, learn/recognize models of other agents, establish/maintain social relationships, use natural cues, exhibit distinctive personality and character, may learn/develop social competencies. It is contestable, however, whether such robots

exist at present. Much current research in social robotics takes place in controlled or contrived situations [18], and/or in robots augmented with pseudo-capabilities through Wizard-of-Oz scenarios [1]. This perhaps impedes a measured understanding of the state of the art in social robotics, because the general public has skewed expectations and perceptions of a given robot's competencies. For example, pet robots in [4] were introduced to a group of participants in an "adopt-a-robot" scenario. Researchers studied human attachment to the robots over an extended period of time, following multiple upgrades to the robot that displayed increased levels of social interactivity. Contrary to expectation, improved sociality showed no significant improvement in attachment at the end of the longitudinal study, as compared to the attachment developed at the onset of the experiment. This example readily illustrates the limits of current social robots' ability to hold interest for prolonged periods, after the initial excitement of novelty has waned. The experimental context also plays a significant role in the perceived 'success' of the interaction [2, 19]. For example, several studies have shown to improve the quality of life of the elderly using the interactive therapeutic robot Paro [20]. However, the context in which the robot was presented to the participants and the effects of scaffolding provided by human mediators in these experiments has not been explicitly studied [19]. Thus, the attributed efficacy of such embodied interactions seems to have been predicated upon parameters that may lie beyond the robots' capabilities. Open experimentation emerges as an alternative to traditional lab-based, narrow context studies [3, 6].

A. Open Experimentation

Open experimentation, or 'in the wild' studies, refer to experiments where the robots are situated in realistic social settings in 'real' environments. Sabanovic et al. [3] argue that social robots should be observed, objectively and analytically by trained social scientists in real-world environments with untrained interactors. While it is difficult to create a context that is broadly unencumbered, environments such as galleries and museums provide rich sites for open experimentation with untrained interactors. In [21] a novel Robot-Robot Interaction study was constructed in a museum setting where the human observer was prevented from directly interacting with the robot, but instead relegated to the role of an observer of the Robot-Robot Interaction that took place. One could argue that the passive observer fulfills the first requirement of objectivity suggested by Sabanovic et al. and being situated in a real-world environment without a clear context. Silvera-Tawil et al. [6] describe another open experiment situated in a gallery space using a minimally articulated humanoid robot. In their study, data clustering through a machine learning algorithm was used to study the nature of the interaction between the robot and humans. Recently, theatre has emerged as another open experimentation site [5, 22-25] providing varyingly broad contexts to situate HRI experiments that evaluate entertainment factors [24-26] to vicariously experienced care robots [5].

Between, cultural spaces such as theatres, museums, gallery

spaces and more general public venues, we find a rich space to situate HRI experiments that offer more dynamic and perhaps more holistic evaluation compared to laboratory settings.

B. Co-experience

A second aspect of constructing a meaningful HRI experiment is the measures of experimentation. As in [3], one could argue these measures should be constructed objectively and elicited from within situations emergent from realistic interactions between the actors.

There have been several efforts to identify and standardize a set of metrics and measurements in HRI research [18, 27-29]. One approach takes an ontological view [29]. Perhaps ambitious given the current technical state of humanoid and other social robots, this approach is directly inspired by psychology research, proposing six benchmarks for the evaluation of a robot agent based on human psychology: Autonomy, Imitation, Intrinsic Moral Value, Moral Accountability, Privacy, and Reciprocity. However, the authors themselves acknowledge the deficiencies in the efficacy of such benchmarks where the technological realities lag far behind the perceived expectations. Nevertheless, these benchmarks provide a suitable starting point for further exploration.

Bartneck et al. [27] have explicitly explored the difficulties in constructing objective measures in HRI and the difficulties faced by the HRI community in general, which they identify as partly structural. In [27], the authors proposed five concepts for consideration: Anthropomorphism, Animacy, Likeability, Perceived Intelligence and Perceived Safety along with five consistent questionnaires using semantic differential scales. Thus far, the question of a reliable set of benchmarks is not settled and as alluded to by Aly et al. [28] this is in part due to the highly multidisciplinary nature of the research area. In fact, in a study conducted at a social robotics workshop most researchers favored a pragmatic approach to benchmarking HRI experiments. RoboCup@Home was highlighted as an example of a useful site which resembles a 'real world' scenario. Three other important insights emerged from the workshop on benchmarking. The HRI community seems to (1) prefer behavioral measures over questionnaires, neuroscientific methods, and system benchmarks, (2) consider situation awareness as the best measure of performance when applied to humans, and (perhaps paradoxically), (3) prefer that robots be judged by their degree of autonomy. The overall aim of evaluation is to see whether the degree of "appropriateness of interaction" is satisfied.

We tend to agree with these sentiments of the research community and have been exploring how to further develop existing measures that address this ethos and arrive at a more nuanced understanding of social interaction between human and nonhuman agents. ED [9] emerges as a potentially rich framework for evaluating interaction and human perception of social cues in HRI research. The three themes of experience-centered design as applied to understanding human interactions and relationship with technology has been described as:

- holistic approach to experience
- continuous engagement and sense-making
- relational or dialogical approach

An important premise of ED approach is the importance it places on prior experience of the interactor as much as what is experienced at the site. The framework also emphasizes the empathetic relation between the maker (designer/programmer/engineer), participant (study participant /museum visitor/end-user) and the emerging artifact (product/tool/robot). This framework acknowledges a complex of aesthetics and expectation between the developer and the user and the iterative process through which all interactions and experiences emerge.

Forlizzi et al, [30] identify three types of experiences: *Experience*, *An Experience* and *Co-Experience*. Of particular interest to the present study is the creation of co-experience. Co-experience, according to [31] describes experiences with products or tools in terms of how individual experiences emerge and change as they become part of social interaction. Thus, in our approach, the robot becomes a social actor engaged in co-creating social experiences with humans.

III. EXPERIMENT

It was essential that the between-participant experiment be conducted in-the-wild using an autonomous robot that represents the current state of the art in HRI. The first generation, collaborative robot Baxter, developed by Rethink Robotics meets these requirements. Baxter essentially performs a number of general automation tasks safely in close proximity to humans. Unlike previous generation industrial robots, Baxter does not require safety cages that isolate human operators from the robot. For the past year, the Powerhouse Museum has exhibited a Baxter robot playing noughts and crosses across with itself (the left arm competing against the right arm – Fig. 2). We aim to improve upon the current installation by creating an interactive game where museum visitors engage in unsupervised game play with a robot. Furthermore, because we are interested in studying unsupervised physical interactions between robots and humans in close proximity, Baxter's size and stability are preferable to smaller, mobile robot platforms.

A. Noughts and Crosses

Noughts and Crosses (or tic-tac-toe) is played on a 3x3 grid between two players who alternatively draw symbols of noughts (O) and crosses (X), and the winner is the player who manages to place three consecutive marks in a horizontal, vertical, or diagonal grid formation.

Our decision to use noughts and crosses was motivated by an observation made by the senior curator at the Museum of Applied Art and Science (MAAS) in Sydney, Australia [32]. He noted that the current installation of a Baxter robot playing Noughts and Crosses (Fig. 2) lacked excitement compared with an industrial robot previously installed at the museum that engaged visitors indirectly using buttons. In that exhibit,

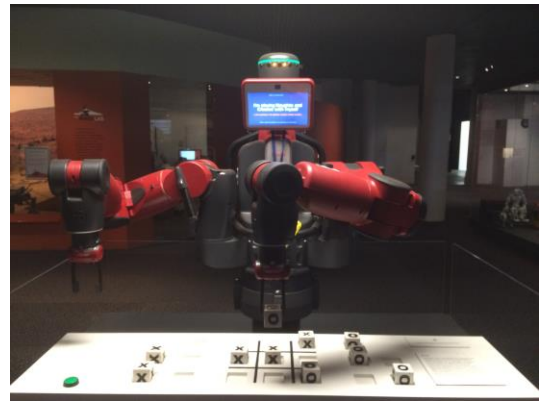


Fig. 2. Baxter robot plays Noughts and Crosses in a fully autonomous, permanent exhibition at Powerhouse Museum (Sydney) where the left arm competes against the right arm.

the robot was situated inside a glass cage and did not interact directly with the visitors. He observed that interest in the Baxter installation was related to the presence of the human programmer mediating the interaction (for example, during the testing phase of the installation).

Noughts and Crosses is also the subject of another important HRI study by Castro-González et al. [18] where participants played games of Tic-Tac-Toe against a Baxter robot that exhibited varying types of motion qualities and explored the effect of a cheating robot. In that study, the researchers confirmed other studies that have shown that movement characteristics influence the human perception of animacy for interactive robots. These lab-based experiments provide the ground truth for our experiments. In addition, game play provides a natural entry point for participants to interact with the robot without being overly conscious of the experiment.

B. Selecting the Venue

The experiments were conducted in a science museum open to the general public. Working in a public space enables us to observe the effects of bystanders, an important social dimension that is difficult to create in a traditional laboratory setting. The selected venue, the National Science and Technology Centre (Questacon) is conducive to studying natural social interaction relative to a laboratory setting. Participation in the study was elective and the subjects were randomized. Because the study participation was voluntary and occurred during regular museum hours, we could not ensure an equal number of participants across all conditions. This is a significant drawback to conducting experimental trials in the wild. Also, the venue enables us to capture a wide demographic; although because of ethics requirements we limited our test-subjects to participants over 16 years old. This is a second, important challenge to conducting experimental trials at science museums, where there are a significant number of children present.

C. Experience-Centered Design Approach

Based on the findings of [15] we reframed the experimental design using an ED framework for measuring co-presence and

co-experience in HCI. This approach enables us to introduce the researcher as an active mediator. In conditions where the researcher was present, we followed a minimal script that used a conversational tone to prompt intuitive and socially-meaningful dialog to emerge. For consistency, the same researcher was involved in the creative design process of the game play (e.g., robot gestures, facial animations, robot arm configurations in pre- post- game play) and related artefacts (e.g., the game board and blocks) and was thoroughly familiar with the set up. What emerges is a rich experimental setup that satisfies the conditions for an experience-centered design. Particularly, the experimental setup is a foundation for a *co-experience*-directed study between the researcher, participant and the robot.

D. Experimental Setup

The Baxter robot has a strong anthropomorphic presence with two articulated arms with 7 degrees of freedom in each arm and a screen display (Fig. 2). The robot was placed at one end of the designated gallery space (Fig. 3 & 4). The screen displayed a minimalist face with a direct gaze (pictured in Fig. 4). A table was placed in front of the robot with an engraved grid and slots for the game board and blocks marked with noughts (O-red) and crosses (X-blue). A computer vision algorithm was used to detect the board and the blocks when placed on the board. We adapted the open source software used in [18] to implement game play. Prior to starting a game, the board must be cleared of any game blocks manually, which was done by a researcher. A new game session begins when a human player places the first block (always a red nought-block) on the board. The robot then makes its first move with a blue cross-block. The game continues until one player wins, or else ends in a draw. This interaction was fully autonomous. The robot only used the right arm to move the blocks from the storage area onto the board. The left arm remained static in a neutral pose throughout the experiment. The robot used an open/close gripper to manipulate the blocks.

A technician stood nearby to prevent any dangerous situations, as the robot was situated in a public setting where small children were present. Children were allowed to participate in game-play, but following ethical guidelines no children were video-recorded, or asked to fill out the survey.

In the first condition, the technician did not interact with the participants, and would clear the board after each game (when

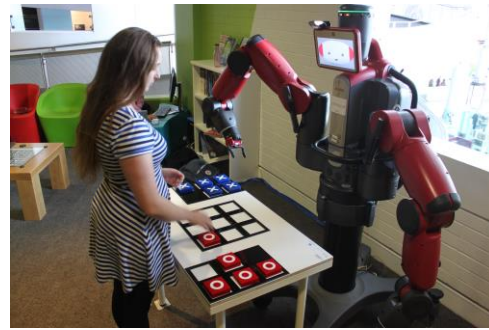


Fig. 4. Baxter Robot engaged in a game of noughts and crosses at Questacon.

participants did not do so), moving any remaining blocks back to the storage area. Three video cameras were used to record the game play: one camera was mounted on the center of the robot at waist height facing the game table and participant. A separate camera mounted on the head of the robot was used for the computer vision algorithm, an off-board camera in the far corner of the room to record the context and a camera close to the robot mounted on a nearby shelf to record the frontal view of the interaction (Fig. 3).

E. Participants

The experiment was conducted over a four-day period. Based on data provided by Questacon [33], we estimate an average of 400 visitors to have seen the Baxter installation each day, but only a fraction of them participated in game play and the experimental study (52 study participants in total). The reason for this smaller sample is that many of the museum visitors were children and therefore ineligible to participate in the study. This is a challenge for working in the wild, particularly in a museum and science centers where families and children constitute a large part of the population. Visitors were allowed to freely interact with the robot in a similar manner to other interactive installations in the gallery. To recruit study participants, a member of the research team would approach museum visitors and onlookers observing a game already in progress (Fig. 4).

When a new participant was selected, they were informed of the general nature of study (“We’re studying human-robot interaction”), the typical duration, and the opportunity to provide feedback through a questionnaire. Participants were also informed of cameras that recorded the interaction and were asked for written consent. Most participants intuitively understood that the robot plays using the right hand and that they should place the first block (O-red). The researcher would inform the participant of the basic setup and then would either walk away from the participant (Condition 1), or would move to the side of the table (Condition 2 and 3) to continue the interaction. At the end of the game, the researcher invited the participant to complete the questionnaire on a tablet. No limits on the number of games a participant could play were imposed. However, none of the participants elected to play a second round. This can probably be attributed to the flow of visitors around the installation, which prompted participants to leave the game after their turn so others could play.

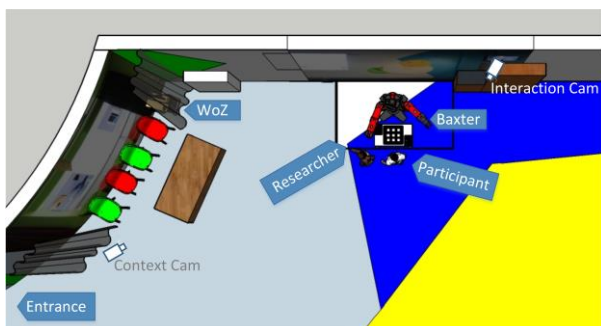


Fig. 3. Baxter Robot Experimental set-up at the Questacon. A Wizard-of-Oz (WoZ) setup was used to turn on the “cheat” function of the robot.

F. The Questionnaire

We wanted to assess the perceived physical presence and feeling of co-presence, and the participant's sense of co-experience and overall impressions during the HRI. We therefore developed a questionnaire that combined questions from Lombard and Ditton [13] (also used in Castro-Gonzalez) for assessing perceived co-presence, and the Nowak and Biocca [14] questionnaire for studying physical presence and perceived social presence for interactions between humans and virtual agents. Our final questionnaire was designed to measure two distinct attributes: co-presence and co-experience. Our reasoning for measuring these two separate questions was to evaluate whether – and to what extent – a robot can contribute to co-experiences. The hallmark of a social robot should not only be the degree of perceived presence or co-presence, but also relates to the robot's ability to generate or contribute to co-experiences.

Co-presence: The co-presence attribute was comprised of the following ten 5-point-scale (Strongly Disagree to Strongly Agree) questions: “The robot was intensely involved in our interaction”, “The robot seemed to find our interaction stimulating”, “The robot created a sense of distance between us”, “The robot seemed detached during our interaction”, “The robot created a sense of closeness between us”, “The robot was interested in interacting with me”, “I did not want a deeper relationship with the robot”, “I wanted to maintain a sense of distance between me and the robot”, “I tried to create a sense of closeness between me and the robot”, and “I was interested in interacting with the robot”.

Experience: The experience attributes consisted of the following thirteen 5-point-scale questions: “How well were you able to observe the body language of the robot? (Not well to Very well)”, “How well were you able to observe the facial expressions of the robot?” (Not well to Very well), “How often did you make a sound (e.g., laugh, speak) in response to the robot? (Never to Always)”, “How often did you smile in response to the robot? (Never to Always)”, “How often did you want to, or did you speak to the researcher? (Never to Always)”, “How often did you want to or did you speak to the robot? (Never to Always)”, “To what extent did you feel mentally immersed in the experience? (Not at all to Very much)”, “How involving was the interaction?” (Not at all to Very much), “How relaxing, or exciting was the experience?”, “How engaging was the game? (Very relaxing to Very exciting)”, “How was the overall movement quality during the interaction? (Very poor to Very good)”, “How comfortable were you interacting with the robot? (Not at all to Very much)”, and “Overall, how satisfying, or enjoyable was the interaction? (Not at all to Very much)”.

The questionnaire also included a field where participants could write general comments about their experience.

G. Conditions

There were three between-subject conditions for the experiment where gameplay occurred: In Condition 1, the physical game board mediates the co-experience. The robot is interactive insofar as it engages in turn-taking that corresponds to formal rules of play. The robot's programmed physical

gestures support real-time game play, and are limited to functional movements. For Condition 2, the researcher mediates the co-experience between the participant and the robot. The robot is interactive as in Condition 1, but in this case the researcher engages in dialogue with both the participant (addressing them by name, encouraging game play) and with the robot (referring to the robot by name e.g., ‘Baxter’, speaking directly to the robot). For Condition 3, the researcher facilitates the interaction as in Condition 2, but the robot is programmed with a “cheat” function, where the robot covers one of the participant's tiles in an effort to cheat win the game. A programmer was required to ‘turn on’ the cheat function prior to game play. A computer was made available for this purpose (marked as WoZ area in Fig. 3).

IV. RESULTS AND ANALYSIS

Fifty-two visitors participated in our experiment; 18 visitors were randomly assigned in Condition 1, 27 visitors were randomly assigned in Condition 2, and 7 visitors were randomly assigned in Condition 3. Table I presents an analysis of the questionnaire including mean values, standard deviations, group means for each attribute, group standard deviations for each attribute, and Cronbach's Alpha values. Table II shows the general comments of the participants relevant to their experience. The Cronbach's Alpha values for all the attributes are above the 0.65 threshold meaning that there is a high degree of internal consistency of the data. All received questionnaires were anonymous.

Fig. 5 presents a boxplot and a bar plot with the ratings for the attributes of co-presence and experience for each condition, and Fig. 6 shows an interaction plot of the group means of each attribute for each condition. A significant condition effect can be observed, as well as a moderate attributes effect, but no interaction effect between conditions and attributes. The mean values for the experience attribute are higher than the mean values for the co-presence attribute, and Condition 2 scores higher than the other two conditions.

We conducted two-way analysis of variance (ANOVA) and the statistical analysis of the results indicated:

- A significant condition effect between all conditions ($F(2; 1148) = 3.2310$; $p = 0.03988$), and a significant attributes effect between all conditions ($F(1; 1148) = 38.0437$; $p = 9.57e-10$).

Further analysis indicated:

- A significant condition effect between Conditions 1 and 3 ($F(1; 552) = 5.037$; $p = 0.0252$), and a significant attributes effect between Conditions 1 and 3 ($F(1; 552) = 20.836$; $p = 6.17e-06$).

- A significant condition effect between Conditions 2 and 3 ($F(1; 740) = 6.027$; $p = 0.0143$), and a significant attributes effect between Conditions 2 and 3 ($F(1; 740) = 20.625$; $p = 6.52e-06$).

- A significant attributes effect between Conditions 1 and 2 ($F(1; 1004) = 35.119$; $p = 4.26e-09$).

TABLE I
ANALYSIS OF THE QUESTIONNAIRE

Attributes	Condition 1 (n:18)			Condition 2 (n:27)			Condition 3 (n:7)		
	Mean	Standard Deviation	Cronbach's Alpha	Mean	Standard Deviation	Cronbach's Alpha	Mean	Standard Deviation	Cronbach's Alpha
I. Co-presence	2.90	0.53	0.65	3.10	0.59	0.71	2.70	0.92	0.81
<i>The robot was intensely involved in our interaction.</i>	3.10	1.00		3.30	1.08		2.42	1.30	
<i>The robot seemed to find our interaction stimulating.</i>	2.50	1.04		3.00	1.26		2.28	1.16	
<i>The robot created a sense of distance between us.</i>	2.90	0.81		3.20	1.08		2.85	0.35	
<i>The robot seemed detached during our interaction.</i>	3.00	0.91		3.10	1.03		3.16	0.37	
<i>The robot created a sense of closeness between us.</i>	2.40	1.00		2.50	1.10		2.33	0.94	
<i>The robot was interested in interacting with me.</i>	2.90	1.22		2.50	1.27		2.14	0.83	
<i>I did not want a deeper relationship with the robot.</i>	3.40	1.14		3.30	1.23		3.14	1.25	
<i>I wanted to maintain a sense of distance between me and the robot.</i>	2.80	1.26		3.70	1.12		2.42	0.50	
<i>I tried to create a sense of closeness between me and the robot.</i>	2.70	1.27		3.00	1.04		2.71	0.70	
<i>I was interested in interacting with the robot.</i>	3.80	0.92		3.70	0.93		3.57	1.17	
II. Experience	3.33	0.65	0.86	3.33	0.71	0.84	3.02	0.98	0.76
<i>How well were you able to observe the body language of the robot?</i>	3.16	1.11		2.73	1.16		2.14	0.83	
<i>How well were you able to observe the facial expressions of the robot?</i>	3.27	1.19		2.46	1.33		2.28	1.40	
<i>How often did you make a sound (e.g., laugh, speak) in response to the robot?</i>	2.94	1.12		2.69	1.26		2.85	1.55	
<i>How often did you smile in response to the robot?</i>	3.44	1.25		3.26	1.19		3.00	0.75	
<i>How often did you want to or did you speak to the researcher?</i>	3.22	0.91		3.80	1.07		3.00	0.90	
<i>How often did you want to or did you speak to the robot?</i>	2.61	1.49		2.76	1.39		2.00	0.90	
<i>To what extent did you feel mentally immersed in the experience?</i>	3.38	0.82		3.50	1.00		2.20	0.74	
<i>How involving was the interaction?</i>	3.50	0.83		3.38	0.92		3.20	1.16	
<i>How relaxing or exciting was the experience?</i>	3.22	0.78		3.5	0.97		3.60	1.01	
<i>How engaging was the game?</i>	3.38	1.00		3.76	1.01		3.60	0.80	
<i>How was the overall movement quality during the interaction?</i>	3.00	1.00		3.26	1.05		3.16	0.68	
<i>How comfortable were you interacting with the robot?</i>	3.88	0.80		3.96	0.80		4.16	1.06	
<i>Overall enjoyment</i>	4.33	0.66		4.33	0.76		4.14	1.12	

TABLE II
GENERAL COMMENTS

Condition	Comments
1	I feel the robot should frown when it loses. Thanks. Danke I wish the robot could talk. I liked playing with the robot because it wasn't too hard. It was very natural.
2	Very involving, look forward to even more interactivity with BAXTER.
3	I love the robot. It was fun playing with the robot playing with my son.

Post hoc analyses were conducted on all possible pairwise contrasts, and more specifically, Tukey HSD tests on 95% family-wise confidence level. The following pairs of groups were found to be statistically significantly different

($p < 0.05$): co-experience and co-presence for Condition 1, co-experience of Condition 2 and co-presence of Condition 1, co-experience of Condition 1 and co-presence of Condition 2, co-experience and co-presence of Condition 2, co-experience of Condition 1 and co-presence of Condition 3, and co-experience of Condition 2 and co-presence of Condition 3.

Not statistically significant results were left out of the analysis.

V. DISCUSSION

We presented an experiment where a Baxter Robot played a game of noughts and crosses (tic-tac-toe) with a participant in a public setting and explored how the perceived co-presence and experience varied with the introduction of a human mediator and cheating behavior by the robot. Based on self-reporting, we found that there were no significant differences in perceived co-presence and experience between participants who interacted with the robot alone, or with an experimenter

as a mediator. However, when the robot cheated (Condition 3), a significant change in perceived co-presence and experience was observed.

The post hoc analyses indicate that co-presence and experience are not synonymous, but rather two distinct categories. This is consistent with ED literature which defines co-experiences as those that typically occur between humans. Objects and tools can have a strong presence without necessarily generating co-experiences. While one might expect a robot with stronger perceived co-presence to be more likely to generate co-experiences, the human perception of a robot's sociality is not only indicated by co-presence but also relates to its ability to generate, or contribute to co-experiences.

Conditions 1 and 2 maintain a separation between experience and co-presence (significant differences), but there is no significant difference between co-presence and experience for Condition 3 (cheating robot). While the number of participants is significantly lower than Conditions 1 and 2, these results might indicate that the cheating robot is perceived more as a social actor and less as an object. Therefore, our future work includes conducting another round of experiments with an equal number of participants across conditions.

It is surprising that there is no significant condition effect between Conditions 1 and 2. We could postulate that for most participants, this was their first encounter with a Baxter robot and in an interactive scenario. Given the novelty of the interaction, as well as the open lab environment where bystanders were often present, the addition of a mediator might not have significantly altered the emerging experience between the participant and the robot. For example, several participants interacted with the robot while their friends watched nearby (Fig. 7). In these instances, the bystanders functioned as mediators and contributors to the co-experience. We are currently analyzing nuanced interactions according to discourse and conversational analysis, which will reveal further insights into the nature of the embodied interactions.

The impact of bystanders and other human mediators has considerable implications for HRI experiment design and argues for the importance of the open lab approach. While further experimentation is needed, our results may indicate the influence of bystanders in similar socially-situated HRI experiments, an essential aspect of most social interactions. However, we must acknowledge that even in the current context of an 'in the wild' experiment, participants were aware that an experiment was being conducted and this awareness might have influenced the reported experience and co-presence ratings.

The significant condition effect between Conditions 1 and 3 as well as between Conditions 2 and 3 follow similar observations described by previous research [18, 34, 35]. As discussed in Castro-González [18], a cheating robot is perceived as less trustworthy by the participants. In our results, this observation perhaps could be interpreted by the lower ratings reported in Condition 3. The cheating robot disrupted the game flow, possibly interrupting the perceived co-presence. While this postulation could account for reduced

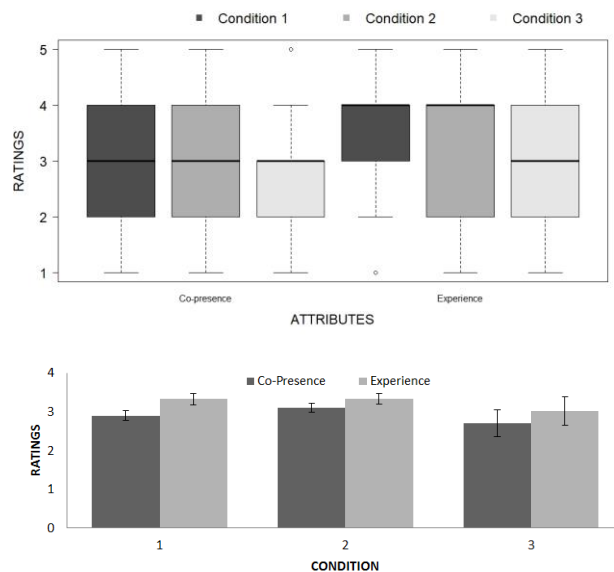


Fig. 5. Boxplot (top) and bar plot showing error bars (bottom) with the ratings for the attributes of co-presence and experience for each condition.

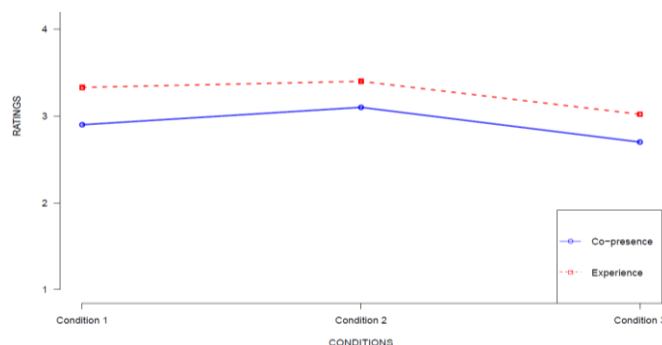


Fig. 6. Interaction plot of the group means of each attribute for each condition.

reporting of co-presence (due to the trust factor), we would expect to see a heightened experience due to the unexpected actions of the robot. The results however, do not indicate a significant interaction effect between the conditions and the attributes.

A. Reliability of Self-reporting

This lack of significant interaction effect between the attributes and conditions seems to stand against what was observed in-situ in real-time, where our observations and subsequent video data analysis revealed different outcomes than those results gleaned from self-reporting alone. Preliminary analysis of video data suggest that participants responded more socially to the robot (looking at the robot more frequently, rather than concentrating attention only on the game board and end effector) in Condition 3 (Fig. 7) than in Conditions 1 and 2. We observed repeatedly that as soon as the robot indicated its intention to cheat (gesture toward placing the block on the participant's tile), participant engagement made a notable change. Participants were more vocal, prone to smiling, and engaged with the robot in socially-significant ways. For instance, in Conditions 1 and 2,

participants were primarily interested in the game and focused their attention primarily on the game board and end effector. Even in Condition 2, the direct interventions by the mediator to engage the participant directly with the robot by looking at the robot's "face" and addressing the robot by name, did little to divert the participant's gaze away from the game board and towards the robot. Given that the other conditions remained equal, video data analysis suggests a significant increase in the social engagement of the participant with the robot from Condition 2 to Condition 3.

There are considerable issues of reliability with self-reporting and HRI experiments [28]. While it is clear from the in-situ observations and the video data that participants had a heightened experience, (for example, the experience attribute of *enjoyment* was clearly more pronounced in Condition 3), this engagement was not reported or reflected in the questionnaire responses. Evaluating and measuring engagement in HRI remains challenging. The discrepancy in what was reported post experience versus what was observed in real-time perhaps could be explored through the dialogical view of experience [9]. One could argue that the spectacle of a cheating robot brings a strong sense of experience as it unfolds in real-time. Here, the experience is centered on three specific centers of value: the participant, the robot and the mediator, and the experience evolves around these three centers spontaneously. As noted by Wright et al. [9], the meaning given to the experience is contingent upon the dialogue that emerges between the three. What is captured in real-time by the camera is the evolving 'dialogue' of the experience. However, the mediator imposes strong negative intonations to the evolving dialogue by reprimanding the robot for 'bad' behavior and correcting the robot by physically maneuvering the misbehaving arm back to its neutral position and 'asking' the robot to drop the offending game block (Fig. 7). Therefore, when the participants retell their experience through the questionnaire, the recollection of the experiences is somewhat skewed by the overall, evolved experience. As Wright et al [9] noted, essentially what is captured by the post-experiment questionnaire is a reconstructed meaning of the original experience contingent on the dialogue that evolved between the three centers of value.

From this line of argument are the questions of (a) efficacy in self-reporting, and (b) the effects of bystanders and mediators on perceived presence and sociality. Our findings highlight the potential bias that might arise in post-experiment self-reporting, when evaluated through the lens of co-experience and in light of other data. A mixed-methods approach that relies on quantitative and qualitative data is one possible way to address this bias to arrive at a more nuanced understanding of relevant factors in human perception of social robots.

VI. CONCLUSIONS AND FUTURE WORK

Further exploration into co-presence and human perception of collaborative robots in museums is warranted. We believe there is strong evidence that suggests a significant discrepancy between self-reporting and real-time interactions. This



Fig. 7. Condition 3. Top-left – The robot is about to cheat. Top-right – robot moves its game block towards the participant's already placed block. Bottom-left – mediator reprimands the robot for 'bad' behavior and moves its arm back to the neutral position. Bottom-right – Robot abandons cheating and get back to playing honest.

discrepancy is compounded by the bystander effects and calls for careful attention when constructing HRI studies to produce meaningful results.

In our experiment, we explored co-presence and experience of collaborative robots through an embodied interaction using a Baxter robot playing a game of noughts and crosses with a human participant in a public setting. We approached the experiment from an Experience Design framework, where the final experiment is an iterative and evolving interaction between the creator, the participant, and the artefact (robot). We propose that this approach is important for creating realistic HRI scenarios, especially when constructing interactive, autonomous robotic experiences in museums and other public environments, where perceived safety and experience are of great importance. This work is an early but significant attempt to develop a framework for museum curators and other early adaptors of socially interactive robots to pursue. Importantly, this work highlights the potential shortcomings of self-reporting and the relevant influences that bystanders and mediators bring to evolving perception of an experience with socially interactive robots.

ACKNOWLEDGMENT

We acknowledge the early contributions made to this project by engineering students from University of Canberra who worked on the software and engineering aspects, and the residencies of graduate students from Aalborg University who contributed to the ED framework for the experiment. We profusely acknowledge the unstinted support provided by the Questacon in carrying out the experiments in the wild.

REFERENCES

- [1] L. Cavedon, C. Kroos, D. Herath, D. Burnham, L. Bishop, Y. Leung, *et al.*, "'C'Mon dude!': Users adapt their behaviour to a robotic agent with an attention model," *International Journal of Human-Computer Studies*, vol. 80, pp. 14-23, 2015/08/01/ 2015.
- [2] D. C. Herath, T. Chapman, A. Tomkins, L. Elliott, M. David, A. Cooper, *et al.*, "A study on wearable robotics - Comfort is in the context," in *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*, 2011, pp. 2969-2974.

- [3] S. Sabanovic, M. P. Michalowski, and R. Simmons, "Robots in the wild: observing human-robot social interaction outside the lab," in *9th IEEE International Workshop on Advanced Motion Control, 2006.*, 2006, pp. 596-601.
- [4] D. C. Herath, C. Kroos, C. Stevens, and D. Burnham, "Adopt-a-robot: A story of attachment (Or the lack thereof)," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2013, pp. 135-136.
- [5] E. Jochum, E. Vlachos, A. Christoffersen, S. G. Nielsen, I. A. Hameed, and Z.-H. Tan, "Using Theatre to Study Interaction with Care Robots," *International Journal of Social Robotics*, vol. 8, pp. 457-470, August 01 2016.
- [6] D. Silvera-Tawil, M. Velonaki, and D. Rye, "Human-robot interaction with humanoid Diamandini using an open experimentation method," in *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2015, pp. 425-430.
- [7] K. Pitsch, T. Dankert, R. Gehle, and S. Wrede, "Referential practices. Effects of a museum guide robot suggesting a deictic repair action to visitors attempting to orient to an exhibit," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 225-231.
- [8] E. Vlachos, E. Jochum, and L.-P. Demers, "The effects of exposure to different social robots on attitudes toward preferences," *Interaction Studies*, vol. 17, pp. 390-404, 2017.
- [9] P. Wright, J. Wallace, and J. McCarthy, "Aesthetics and experience-centered design," *ACM Trans. Comput.-Hum. Interact.*, vol. 15, pp. 1-21, 2008.
- [10] M. Conell, "Machines with attitude," ed, 2014 Available: <https://youtu.be/vzazB8ptMI0?t=1s>.
- [11] E. Short, J. Hart, M. Vu, and B. Scassellati, "No fair!! an interaction with a cheating robot," in *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, 2010, pp. 219-226.
- [12] C. Fitzgerald, "Developing Baxter," in *2013 IEEE Conference on Technologies for Practical Robot Applications (TePRA)*, 2013, pp. 1-6.
- [13] M. Lombard, T. B. Ditton, D. Crane, B. Davis, G. Gil-Egui, K. Horvath, et al., "Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument," in *Third international workshop on presence, delft, the netherlands*, 2000, pp. 2-4.
- [14] K. L. Nowak and F. Biocca, "The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments," *Presence*, vol. 12, pp. 481-494, 2003.
- [15] C. G. Oh and J. Park, "From mechanical metamorphosis to empathic interaction: a historical overview of robotic creatures," *J. Hum.-Robot Interact.*, vol. 3, pp. 4-19, 2014.
- [16] C. Breazeal, "Toward sociable robots," *Robotics and Autonomous Systems*, vol. 42, pp. 167-175, 2003/03/31/ 2003.
- [17] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, pp. 143-166, 2003/03/31/ 2003.
- [18] Á. Castro-González, H. Admoni, and B. Scassellati, "Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness and unpleasantness," *International Journal of Human-Computer Studies*, vol. 90, pp. 27-38, 2016/06/01/ 2016.
- [19] M. Birks, M. Bodak, J. Barlas, J. Harwood, and M. Pether, "Robotic Seals as Therapeutic Tools in an Aged Care Facility: A Qualitative Study," *Journal of Aging Research*, vol. 2016, p. 8569602, 11/20 2016.
- [20] K. Wada, T. Shibata, T. Saito, K. Sakamoto, and K. Tanie, "Psychological and Social Effects of One Year Robot Assisted Activity on Elderly People at a Health Service Facility for the Aged," in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 2005, pp. 2785-2790.
- [21] C. Kroos and D. C. Herath, "We, Robots: Correlated Behaviour as Observed by Humans," in *Social Robotics: 6th International Conference, ICSR 2014, Sydney, NSW, Australia, October 27-29, 2014. Proceedings*, M. Beetz, B. Johnston, and M.-A. Williams, Eds., ed Cham: Springer International Publishing, 2014, pp. 229-238.
- [22] W. D. Smart, A. Pileggi, and L. Takayama, "HRI 2010 workshop 1: What do collaborations with the arts have to say about HRI?," in *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, 2010, pp. 3-3.
- [23] H. Knight, "Eight Lessons Learned about Non-verbal Interactions through Robot Theater," in *Social Robotics: Third International Conference, ICSR 2011, Amsterdam, The Netherlands, November 24-25, 2011. Proceedings*, B. Mutlu, C. Bartneck, J. Ham, V. Evers, and T. Kanda, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 42-51.
- [24] N. Mavridis and D. Hanson, "The IbnSina Center: An augmented reality theater with intelligent robotic and virtual characters," in *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009, pp. 681-686.
- [25] K. Ogawa, K. Taura, and H. Ishiguro, "Possibilities of Androids as poetry-reciting agent," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012, pp. 565-570.
- [26] E. Jochum, P. Millar, and D. Nuñez, "Sequence and chance: Design and control methods for entertainment robots," *Robotics and Autonomous Systems*, vol. 87, pp. 372-380, 2017/01/01/ 2017.
- [27] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots," *International Journal of Social Robotics*, vol. 1, pp. 71-81, January 01 2009.
- [28] A. Aly, S. Griffiths, and F. Stramandinoli, "Metrics and benchmarks in human-robot interaction: Recent advances in cognitive robotics," *Cognitive Systems Research*, vol. 43, pp. 313-323, 2017/06/01/ 2017.
- [29] P. H. Kahn, H. Ishiguro, B. Friedman, and T. Kanda, "What is a Human? - Toward Psychological Benchmarks in the Field of Human-Robot Interaction," in *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*, 2006, pp. 364-371.
- [30] J. Forlizzi and K. Battarbee, "Understanding experience in interactive systems," presented at the Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques, Cambridge, MA, USA, 2004.
- [31] K. Battarbee and I. Koskinen, "Co-experience: user experience as interaction," *CoDesign*, vol. 1, pp. 5-18, 2005/03/01 2005.
- [32] M. Cornell, "Baxter Installation at the Powerhouse Museum," D. C. Herath, Ed., ed, 2016.
- [33] J. Wilkins, "Questacon Visitor Figures (May 2015 and 2016)," C. H. Damith, Ed., ed, 2017.
- [34] J. Young, J. Sung, A. Vaida, E. Sharlin, T. Igarashi, H. Christensen, et al. (2011), Evaluating Human-Robot Interaction. *International Journal of Social Robotics 3(1)*, 53-67.
- [35] V. Groom, J. Chen, T. Johnson, F. A. Kara, and C. Nass, "Critic, compatriot, or chump?: responses to robot blame attribution," presented at the Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction, Osaka, Japan, 2010.



Damith C. Herath (M'02) Dr. Herath holds a PhD in Robotics from the University of Technology, Sydney and BSc (Hons) degree in Production Engineering from the University of Peradeniya, Sri Lanka. He is currently an Assistant Professor in Software Engineering attached to the Human-Centered Technology Research Centre at the University of Canberra. Damith is a multi-award winning entrepreneur.



Elizabeth Jochum Dr. Jochum received a B.A. from Wellesley College and an M.A. and Ph.D. in theatre from the University of Colorado, Boulder.
In 2014, she was a postdoctoral researcher with the Geminoid DK lab at

Aalborg University (Denmark). Since 2015, she has been an Assistant Professor in the Research Laboratory for Art and Technology (RELATE) within the Department of Communication and Psychology at Aalborg University. Her research focuses on the intersection of robotics, art, and performance. She is the co-founder of the Robot Culture and Aesthetics (ROCA) at the University of Copenhagen and a member of Aalborg U Robotics, and serves on the editorial board of Global Performance Studies. Her publications appear in *Robots and Art*, *Controls and Art*, *Robotics and Autonomous Systems*, and the *International Journal of Social Robotics*.



Evgenios Vlachos received the M.S. degree in electronic automation from National and Kapodistrian University of Athens, Greece in 2009, and the M.A. and Ph.D, degrees in human centered communication and informatics from Aalborg University, Denmark in 2012 and 2016 respectively. From 2011 to 2015, he participated in the Geminoid-DK project. Since 2016, he was a Postdoctoral Researcher with the Electronic Systems Department, Aalborg University, Denmark employed at the project “Durable Interaction with Socially Intelligent Robots” supported by the Danish Council for Independent Research. Currently, he is a Research Librarian at the University of Southern Denmark. Dr. Vlachos’ awards and honors include the Best Interactive Presentation Award at Int. Conference on Social Robotics in 2012, and three annual scholarships from the Greek State Scholarship Foundation (IKY).