Do you want to make your robot warmer? Make it more reactive!

Alexa Bertó Giménez, Enrique Fernández-Rodicio, Álvaro Castro-González, Miguel A. Salichs

***Abstract*—Endowing robots with the ability to respond appropriately to stimuli contributes to the perception of an illusion of "life" in robots, which is determinant for their acceptance as companions. This work aims to study how a series of bio-inspired reactive responses impact on the way in which participants perceive a social robot. In particular, the proposed system endows the robot with the ability to react to stimuli that are not only related to the current task but are also related to other external events. We conducted an experiment where the participants observed a video-recorded interaction with two robots: one was able to respond to both task-related and non-task-related events, while the other was only able to react to task-related events. To evaluate the experiment, we used the RoSAS questionnaire. The results yielded significant differences for two factors, showing that the addition of responses to non-task-related stimuli increased the robot's warmth and competence.**

***Index Terms*—reactive robots, reactive robots, human-robot interaction, social robotics, social robots, liveliness.**

I. Introduction

It is only a matter of time before machines are no longer perceived as mere tools but as *agents* who live together with us and help us in many different daily activities (e.g. [1] [2]). This is particularly relevant for social robots because they are machines that are intended to coexist and interact socially with humans.

For social robots to be perceived as companions instead of just tools, they need to be endowed with a series of functionalities. These features allow them to interact with humans in a similar way that humans interact with each other [3]. However, the first step to establish these interactions is to be able to identify the interlocutor as a valid social actor [4]. While works such as the Media Equation Theory [5] or the work of Nass et al. [6] show that people tend to interact with media, like computers, as if they were social actors without the need for this media to be identified as human or human-like, there is another body of research that shows the benefits of having a lively appearance for interactions. For example, the work of Csibra et al. [7] shows that endowing non-human entities with the ability to move on their own helps people to attribute these entities with intentions, beliefs, desires, and purposes. Another example is the Biophilia hypothesis, which proposes that humans have a tendency to affiliate with life. Originally proposed by Erich Fromm [8], over the years this theory has attracted both defenders [9], [10] and detractors [11]. Based on the evidence, it seems that humans might tend to feel affiliations towards life forms and that, even if its not essential, giving a robot an appearance of animacy might help it to create relationships with its users.

The term animation is understood as "giving life" to an object that is inanimate; that is, giving it the ability to move and behave like an animate being full of energy [12] [13]. Many robot movements and behaviours are inspired by humans or animals. Observing and studying the behaviours of living beings can give an idea of how a robot can transmit different messages thanks to its physical characteristics and capabilities [14]. Based on the literature, we believe that the robot's ability to interact socially may be aided by the use of lively behaviour imitation [15].

Animation depends not only on the robot's size or appearance but also on the inclusion of lively behaviours, such as the ability to respond to stimuli or the adaptation of the robot's responses [16]. This study asks the following research question: What kind of stimuli, or events, should our robots be aware of, and consequently react to? We hypothesise that a robot that reacts not only to the events directly related to its current task but also to other events will be better perceived by the users.

We aim to evaluate the importance of endowing social robots with reactive behaviours, which respond to external stimuli immediately and without losing the context of the interaction. According to Sowmya, reactive robots[1] exhibit behaviour as reactions to events [17]. These reactions can be of a very different nature, some closer to the reflexes that can be observed in animals (e.g., changing the body position to maintain balance [18]) and others can be more abstract (e.g., emotionally reacting to the content of a conversation [19]).

Although reactiveness is a term that might involve a wide range of behaviors and situations, we have established a distinction between two types of reactions: task-related and non-task-related. Task-related responses are reactions to situations

Authors are with the Robotics Lab, Department of Systems Engineering and Automation, Universidad Carlos III de Madrid, España. E-mails: alexa.berto@alumnos.uc3m.com, enrifern@ing.uc3m.es, acgonzal@ing.uc3m.es, salichs@ing.uc3m.es

The research leading to these results has received funding from the projects: Robots Sociales para Estimulación Física, Cognitiva y Afectiva de Mayores (RoSEs), RTI2018-096338-B-I00, funded by the Ministerio de Ciencia, Innovación y Universidades, Spanish Goverment; Robots sociales para mitigar la soledad y el aislamiento en mayores (SoRoLI), PID2021-123941OA-I00, funded by Agencia Estatal de Investigación, Ministerio de Ciencia e Innovación, Spanish Goverment; and the Multiannual Agreement with UC3M ("Fostering Young Doctors Research", SMM4HRI-CM-UC3M) in the context of the V PRICIT (Research and Technological Innovation Regional Programme) by the Madrid Government (Comunidad de Madrid-Spain).

[1]Although some authors refer to these robots as *responsive* robots (see Section II), for the sake of clarity, we have used the term *reactive* robots most of the time. In our analysis of the state of the art, Section II, we have maintained the original terms used by the authors.

that arise in the context, and as a result of the interaction or activity that the robot is currently performing. An example would be the ability to properly acknowledge when the user makes a good play in a game or the ability to react appropriately when the user discloses personal information. Non-task related responses are reactions to events that take place in the environment where the robot is placed but are not directly tied to the interaction or activity at hand. An example would be reacting to a loud noise while playing a game with the user. Ignoring that event would not hinder the interaction but it could affect to how the robot is perceived by the user because this is something that humans are innately capable of doing.

This study will focus on how the user's perceptions of a robot change when these non-task-related reactions are integrated in the robot. The design of the robot's reactions has taken inspiration from human behaviours because we believe that it is not only important that the robot is able to react to the stimuli considered but that the reaction feels appropriate to the user interacting with the robot.

The rest of the paper is structured as follows. In Section II, related works in the area of robot reactiveness are reviewed. Section III introduces the behaviours that have been developed. In Section IV, we present the experiment, which was designed to evaluate the behaviours through the RoSAS questionnaire [20]. The results are presented in Section V. In Section VI, we discuss the results and we state some of the limitations that we observed during the development of the system and its evaluation. Finally, our conclusions are given in Section VII.

II. State of the art

In this section, we will review the literature of reactiveness in robotics. In particular, we focus on studies that have analysed the effect that the addition of reactive behaviours has had on how the users perceived the robot.

In 2016, Cameron et al. [21] studied the influence that the autonomy and responsiveness of a robot had on children's behaviours. They recorded an interaction with the Zeno robot, where the children had to complete a series of physical activities, such as jogging in place. In this experiment, the authors developed two conditions: 'solo interaction', with a fully-autonomous and responsive robot; and 'adult-assisted interaction', requiring adult assistance to interact with the robot because the robot was only adult speech responsive. The recordings were analysed to count the instances that the children looked towards the robot or the researcher during the interaction. When looking at the robot, the authors considered two situations: anticipatory looks, where the children looked at the robot before it started speaking; and reactive looks, where the child looked at the robot after it started speaking. Then, through questionnaires such as the Self-Assessment Manikin (SAM), the authors obtained measures of task engagement, human emotional facial expression, gaze direction and reported affect. The participants' facial expressions were used to detect the following expressions: neutral, happy, sad, angry, surprised, scared, and disgusted.

The researchers found that participants in the solo interaction condition completed significantly more physical activity in comparison to those in the adult-assisted condition. They also looked towards the robot in anticipation on more occasions and for a greater percent of occasions. In contrast, participants looked toward the robot in reaction on fewer occasions and for a smaller percent of occasions. In the solo condition, they also showed greater average sadness than in the adult-assisted condition (no significant differences were found for other expressions). Researchers found greater average valence (measured with the SAM) in the adult-assisted than in the solo condition. However, there were no significant differences in the ratings of enjoyment of the interaction, and neither for arousal nor dominance.

In 2014, Hoffman et al. [22] studied the effects of a robot being perceived as responsive by a human. They used the robot Travis and implemented non-verbal gestures, such as nodding or looking away and textual responses. They then studied a person's perception of the robot's character traits and attractiveness. The participants had to tell the robot a story, and then the robot tried to respond with a positive or negative attitude. To respond with a positive attitude, the robot displayed a subtle animacy: a gentle sway and short affirmative nods in response to human speech. To show a negative attitude, the robot displayed decreased animacy: no confirmation gestures and occasionally looking away.

Under the positive responsive robot condition, the participants rated the robot as being more responsive and having higher social traits. It was also perceived as more attractive when it reacted positively. No significant differences were found in terms of competence.

A year later, Hoffman et al. [19] presented Kip1, a companion robot that was designed to promote non-aggressive conversation between people. They wanted to increase people's awareness of the effect of their behaviour towards others during a conversation of a topic that they disagreed on. To achieve this, Kip1 was able to react to the conversation that was happening in front of it. The authors tested the effect that these reactions had over the participants through an experiment where two different conditions were compared. In the experiment condition, the participating couples shared a room with a robot endowed with the ability to react to their speech using physical gestures. The robot can be either curious towards the conversation or scared by it. Curiosity translates in the robot stretching towards the speaker, looking around, and raising its head. Fear translates in a retracted posture, where the robot cowers and shivers. In the control condition the robot was animated but not reactive, it just displayed a regular breathing pattern.

This study revealed that couples in the experiment condition referred to the robot slightly more often than those in the control condition. This means that when the robot reacted to their conversation, they paid more attention. They also rated the robot's social human traits as being higher and slightly more similar to them.

In 2020, Rifiniski et al. [23] presented a study that extended the work of Hoffman et al. [19]. They also used Kip1 for their study with three between-subjects conditions, with gestures such as gaze gesture, leaning gesture or breath-like gesture. The robot could be 'Responsive to Speaker', 'Responsive to Addressee' or 'Non-responsive'. Under the 'Responsive to Speaker' and 'Responsive to Addressee' conditions, its behaviour was a combination of the previous gestures. Under the 'Non-responsive' condition, the robot just seemed like it was 'breathing'. Participants in the experiment took part in a debate where the robot was present. They were then interviewed and completed a questionnaire.

The responsive conditions resulted in higher conversation-partner liking ratings than the non-responsive condition. The participants that evaluated the robot under the responsive conditions also rated animacy and likability of the robot more highly. Under the 'Responsive to Speaker' condition, Kip1 was perceived as having a higher intelligence than under the baseline condition, which resulted in more smiles received in comparison to the 'Responsive to Addressee' condition and the baseline condition.

In 2016, Hoffman et al. [24] used the robot Travis to study human–robot interaction in relation to media consumption, nonverbal behavior, timing, and physical presence.

In a first study, Travis moved constantly to the rhythm of the beat under the On Beat condition. Under the Off Beat condition, it moved at the same tempo as the music but off beat. Finally, under the Static condition, the robot remained static. Each participant was assigned one condition and listened to three songs with the robot as the "speaker device". After each song, they had to answer three questions measuring song liking. Song liking for the On Beat condition was the highest, the lowest for Static, and for Off Beat in-between. In addition, positive human traits were attributed to the robot under the On Beat and Off Beat conditions (these attributes are a composite measure of friendliness, confidence, warmth, cooperativeness, and sociability), and participants rated the robot as more human-like. They concluded that people like songs better with the robot responding to the music and that such a robot is rated more highly on positive human character traits.

In a second study, the participants watched a clip from a comedy show alone under the Control condition or with Travis next to them being responsive to the clip under the Responsive condition. The robot's gestures were synchronised with the video's laughs. It nodded, looked back-and-forth between the screen and the participant, and also used other head movements. Finally, under the Present condition, Travis was present but not reacting to the events. The participants then filled out an enjoyment questionnaire and were asked about their music listening habits, solitary or social (between-subject variable). The authors found a significant difference for Viewing Habits among Social participants under both the Responsive and Present conditions compared to the Control condition, but not for Solitary participants between the three conditions. Finally, under the Responding condition, the participants attributed more positive characteristics to Travis than under the Present condition.

That same year, Birnbaum et al. [25] also used the robot Travis for an experiment where the participants were led to believe that they were testing an algorithm that allows robots to understand speech, and that Travis would try to respond. Under the Responsive condition, Travis reacted with gestures and written text on its screen. Under the Control condition, Travis did not respond during the experiment with gestures, just with text. Travis was perceived as being more responsive, social and competent under the Responsive condition than under the Control condition, but it was not perceived as more attractive. They also found that participants were more interested in robot companionship when the robot was responsive because they exhibited more approach behaviours (e.g., physical proximity, leaning toward the robot, eye contact and smiling).

Similar to the works presented above, we also seek to evaluate the effect that a reactive robot could have on how it is perceived by the users. However, while the previous works focused on task-related reactions (i.e., reactions to events related to the task the robot is actively engaged in), we study how the integration of reactions to stimuli not connected to the task at hand in a robot affects how the robot is perceived by the users. In addition, while the previous studies mainly focused on non-verbal behaviours or text on a screen, we tried to imitate the human reaction to certain stimuli by creating multimodal human-like behaviours. We achieved this aim thanks to the appearance of our robot, which bears more resemblance to a human body than those considered in previous works because it has arms, eyes, voice and even a heart. Another difference with the works reviewed in this section is the use of the RoSAS questionnaire [20], which is an 18-item scale, to measure people's perceptions of the social attributes of a robot. Thanks to the RoSAS, we consider not just positive aspects (e.g., the interaction, emotion, or capability of the robot) but also negative aspects (e.g., aggressiveness).

III. Integration of non-task-related reactive behaviours

In this work, we analyse the impact that non-task-related reactiveness in robots (i.e., robots that react to unexpected events) has on their attributions. We expect that including behaviours as a response to these events will make the robot look warmer and more competent, while lowering (or at least not raising) the feeling of discomfort felt by the users.

A series of multimodal gestures associated with certain events, which can be triggered during human-robot interactions, have been implemented. When we designed these behaviours, we drew inspiration from human responses to these same events. Thanks to the robot's perceptual capabilities and the gestures developed, the robot is able to react to different stimuli coming from the environment or triggered by a change in its internal state. When a stimuli is received, the system evaluates its type and chooses the response that fits best. This response is performed with a combination of multimodal actions involving different expressive capabilities (e.g., movements, coloured lights, or voice). The stimuli and their associated responses are shown in Table I. Some examples can be seen in Figure 1.

For stimuli such as the user touching the robot, the robot reacts by looking towards the part of its body that has been touched. When the stimulus is a caress, a slap, or a tickle, the robot combines the corresponding non-verbal gesture with an

TABLE I
STIMULI THAT THE ROBOT CAN REACT TO AND ITS DESCRIPTION.

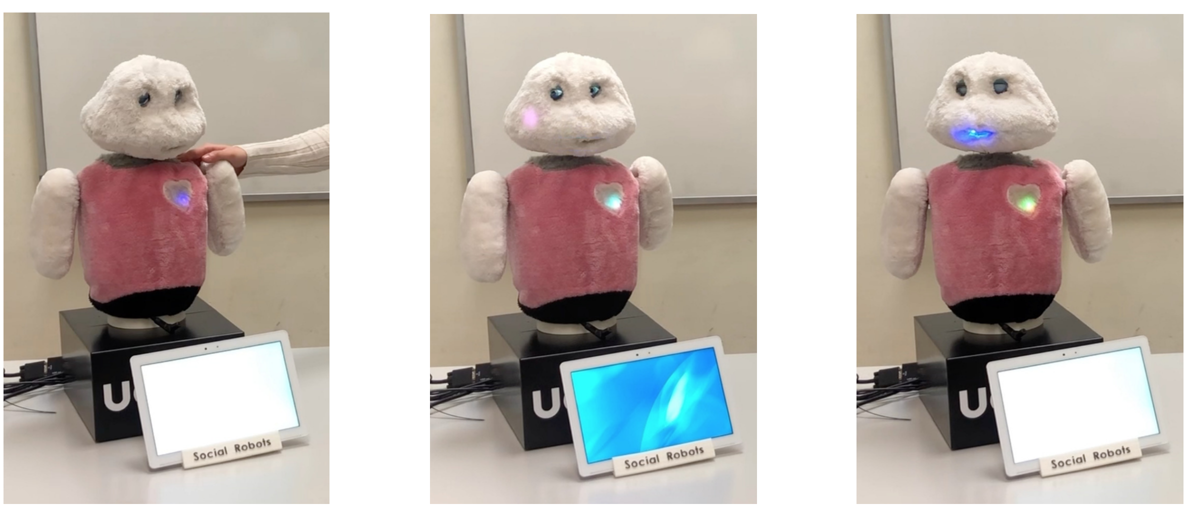
Stimuli	**Description of gesture**
The user touches the robot's head	Head upwards, eyes looking up, arms slightly raised
The user touches the robot's belly	Head downwards, eyes looking down, arms slightly backwards
The user touches the robot's right shoulder	Head and neck turned towards right shoulder, eyes looking down to the right, base slightly turned to the right
The user touches the robot's left shoulder	Head and neck turned to left shoulder, eyes looking down to the left, base slightly turned to the left
The user tickles robot	Head upwards and neck turning alternately right and left, base rotating alternately right and left according to neck, eyes with a joyful expression, green heart, laugh and sentence "You tickle me"
The user caresses the robot	Head first downwards and then upwards and neck turned to the left, right arm to the front and left arm to the rear, base slightly to the left, eyes with a joyful expression, green heart, rosy cheeks, sentence "Hum, I love to be caressed"
The user slaps the robot	Head downwards and neck turning alternately to the right and left, arms forwards and backwards alternating the right and the left one, base rotating alternately right and left according to neck, eyes with an angry expression, red heart, sentence "You are hurting me"
The robot detects high core temperature	Head upwards and neck turning alternately to the right and left, arms raised and moving slightly up and down, alternating the right and the left one, base rotating alternately right and left according to neck, eyes with surprised expression, yellow heart, sentence "Oh, my engines are heating up"
The robot detects a linear movement	Head slightly upwards, eyes with a joyful expression, green heart, sentence "Where are we going?"
The robot detects an angular change	Head upwards, arms slightly raised, eyes with surprised expression, sentence "Be careful, I don't want to fall down"
The robot hears a loud sound	Eyes with surprised expression, head and base turned towards the sound, sentence "Uh, I got scared"

utterance. We also implemented reactions to changes in the robot's internal state, such as the robot expressing its concern when its internal temperature is high, or when the orientation of the robot changes and it could be falling. In addition, the robot reacts to stimuli coming from the environment, such as a loud sound, showing surprise or fear.

This system has been design to be compatible with other activities and interrupt them if necessary. Given that the reactions that we have considered have to be immediate, if any of the stimuli described in Table I is detected, then the robot pauses the current interaction and reacts to the input received. After completing the reactive behaviour, it retakes the previous interaction. Because these reactions are brief, we expect that this interruption will not hinder the interaction.

IV. EVALUATION OF THE SYSTEM

To evaluate the impact of the non-task-related reactive behaviours, we have conducted an evaluation to determine if these reactive behaviours improved the users' perceptions of the robot. Due to the COVID-19 pandemic and the restrictions on face-to-face evaluations, we proceeded with a video-based evaluation using Mini robots (see Figure 1). Mini is a 50-cm high desktop robot whose purpose is to support the elderly and caregivers in their daily life [26]. With a cartoon-like the



(a) The robot looks at the point of contact after being touched by the user

(b) The robot reacts to being caressed by the user

(c) The robot reacts to being tickled by the user

Fig. 1. Examples of the robot Mini's reactions after different stimuli

appearance, it has 5 degrees of freedom, expressive eyes and an LED-based beating heart.

Video-based evaluations make it easier to reach a larger population. This implies a broader range of age and educational backgrounds, and also a larger number of observers, which is important for statistical analysis. Moreover, this is a controlled environment, so the conditions and stimuli the participants watch are the same [27].

A. Description of the scenario

The interaction that was shown in the video consisted of a participant playing with two Mini robots and an instructor guiding the game. The robots played a quiz game, where they asked questions in turn and the participant answered them.

For the recording, two robots were included in the scenario, as shown in Figure 2. They were placed on a table with their backs to the instructor. The instructor was standing and directing the situation, showing the robots and the user how to proceed. The user sat in front of the two robots and interacted with the robots during the recording. Finally, behind the user, there was the camera person and a camera attached to a tripod.

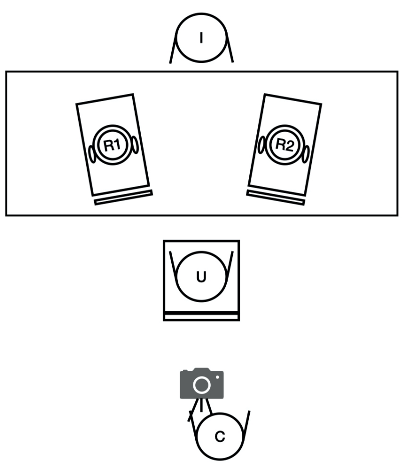


Fig. 2. Scheme of the scenario. I: instructor, R1: robot1, R2: robot2, U: user, and C: cameraman.

B. Conditions

For this experiment, we opted for a simultaneous within-subjects evaluation, with two conditions at the same time. Under the Task and Non Task-Related reactive condition (TNTR condition from now on), the robot was able to react to both task-related and non-task-related events. The list of non-task-related events that the robot can react to is given in Table I. Regarding the task-related events, this includes actions such as acknowledging the user's answers during the quiz game and giving appropriate feedback (e.g., congratulating the user if the answer is correct). Under the Task-Related reactive condition (TR condition from now on), the robot was able to react exclusively to task-related events, ignoring those stimuli that are not connected to the task in hand. From here on, the robot under the TNTR condition will be referred to as the TNTR robot, and the robot under the TR condition will be referred to as the TR robot.

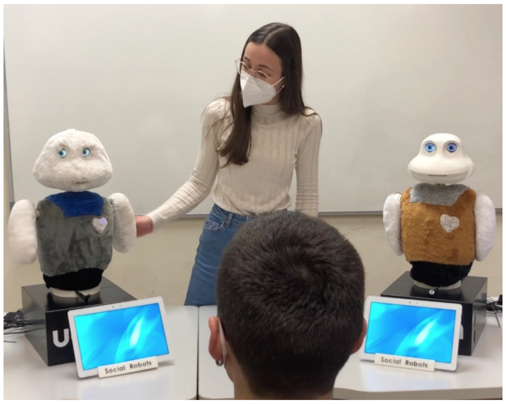
Within-subjects experiments have the advantage of not being dependant on the random assignment of participants to the different conditions and of reducing the number of participants required for the experiment. However, they are also less conservative and can introduce bias in the participants depending on the order in which the conditions are presented [28]. In particular, we opted for a simultaneous within subjects experiment [29]; where instead of testing one condition at a time, the participants are asked to evaluate both conditions simultaneously. Hsee et al. [30] observed a reversal on the judgement that people make of different options depending on whether or not both options are evaluated at the same time. They proposed the evaluability hypothesis, which posits that some attributes are harder to evaluate than others and that these attributes can have a bigger effect when performing a joint evaluation. In our case, both robots are able to appropriately react to events related to the task in hand (i.e., they are both able to conduct a game with the user, to process his/her answers correctly, to provide appropriate feedback, etc.). Because of this, we believe that it is possible that the TR robot is already perceived as warm and competent. As a solution, we opted for presenting both robots in the same video, so the differences were perceived better. However, we avoided telling the participants about the reactive behaviours and instead led them to believe that they would be evaluating the interaction in general. This way, we wanted participants to avoid focusing their attention on the reactive behaviours.

To avoid the bias that might appear due to the appearance of the robots, two videos were recorded in which the colour of the robot's body was changed. In the first video, as shown in Figure 3a, the TNTR robot is grey and blue, and the TR robot has a brown and grey body. In the second video, as shown in Figure 3b, the TNTR robot is pink, and the TR robot has a grey and blue body.

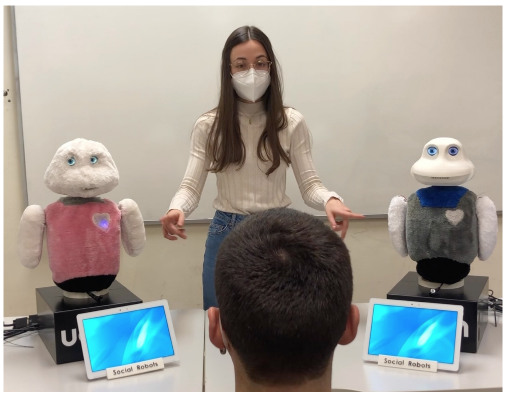
C. Description of the interaction

The videos are 4 and a half minutes long, and in both cases the interaction depicted was exactly the same, as follows [2] (Figure 4 shows the timeline of the interactions). The instructor begins by asking the user if they have ever played with robots, and the answer is negative. The TR robot is then introduced and placed in the scenario. While explaining what it can do, the instructor places their hand on the robot's right shoulder. During this presentation, the softness of the robot is mentioned, so the user is invited to touch it and caress its belly. Then, the instructor introduces the TNTR robot into the scenario. It reacts in surprise and says: "Be careful, I don't want to fall down." It is explained the TR robot does exactly the same, while the instructor also places their hand on its right shoulder. In response, the robot looks towards the shoulder that the instructor is touching. The instructor finishes

[2]Notice that the robot's utterances presented here have been translated from common Spanish sentences that are used as reactions to the considered stimuli. Due to the translation, the affective load of these sentences might be slightly different.



(a) Staging of the first video



(b) Staging of the second video

Fig. 3. Stages of the video-based evaluation

the explanation and again invites the user to caress the robot. The TNTR robot laughs and says: "I love to be caressed."

The instructor then explains to the user that they are going to proceed to play with the robots. The first game consists of four images, where one image differs to the others, so the user has to tell the robot which one is the odd one out. The games have been introduced to make the user think that we are evaluating the games instead of the reactiveness of the robot, trying to remove any bias on their part. During the games, the robots randomly move their arms, base and neck. The user starts playing with the TR robot, and chooses the image that does not fit among the four shown. They then play with the TNTR robot in the same game, and in the middle of the interaction they hear a thump. The TR robot does not show any acknowledgment of the noise, while the TNTR robot responds by simulating that it is startled. It turns to where the noise comes from and says: "I got scared." The interaction continues, and then the user and the TNTR robot finish the game.

The instructor then indicates that they are going to play the sayings game, this time starting with the TNTR robot. The robot displays jumbled words of a saying on its tablet, and the user reconstructs it to form a popular saying. The user then plays with the TR robot. It says the beginning of a saying and the user finishes the saying by voice.

After completing the interaction, the instructor congratulates the two robots on their good work while tickling their heads. Because of this, the TNTR robot detects the tickling and reacts by laughing and saying: "You tickle me!" Meanwhile, the TR robot remains still. The instructor places their hand on the left shoulder of both robots, and again the TNTR robot turns towards that shoulder. Meanwhile, she asks the user if they have any questions about the robots. The user asks if they have more games, and while the instructor explains everything that the robots can do, the TNTR robot starts to raise its arms and says: "My engines are warming up." The instructor explains that they have to rest and thanks the user for their participation.

In the videos, which robot starts to perform each of the two games has been randomised, alternating between TNTR and TR, with the aim of reducing bias. Regardless, the reactions of each robot are always the same as described above.

To provide a clear understanding of the experiment, both videos are available: (i) one where the TNTR robot is grey[3] and (ii) other where the TNTR robot is pink[4].

D. Questionnaire

The evaluation was conducted using an online questionnaire that was developed through Google Forms. On the front page, before entering any personal data, the participants were given a description of the evaluation (i.e., that they will take part on a social robotics experiment where they will be asked to watch a video and respond a series of questions), their rights regarding data privacy, and the consent form that they had to accept before starting the questionnaire. The Data Protection Office from Universidad Carlos III de Madrid approved the evaluation.

The next section asks for a series of demographic data, which will then be used to draw conclusions based on different population groups, as follows: gender, age, educational background and if they live alone or not. Then, in the following section of the questionnaire, a video is shown and the participant is invited to watch it completely. After showing the video, the questionnaire includes a control question to check if the participant has viewed the entire video. This question asked the participant to give the name of the two robots in an image extracted from the video. This also allowed us to check if the participants could actually tell the robots apart. The next section of the questionnaire presents a series of general questions about the interaction the participant witnessed, as follows: (i) Which of the two robots do you prefer to play with?; (ii) Which of the two robots would you have at home?; (iii) Which robot would you like to keep you company?; and (iv) Which robot do you feel more comfortable with? The participants are asked to justify their answers, which enabled us to detect cases in which they may have answered randomly. Finally, the participant is asked: Do you think that the robots' reactions are natural (noises, caresses, etc.)? And again, they are asked to give a justification for their answer, which in this case is also a way of obtaining feedback on the programmed behaviours.

The next section of the questionnaire is used to assess the robots in the video, through the use of the RoSAS ques-

[3]https://youtu.be/cGBQPIBGJTw
[4]https://youtu.be/pnVi15QQRfY

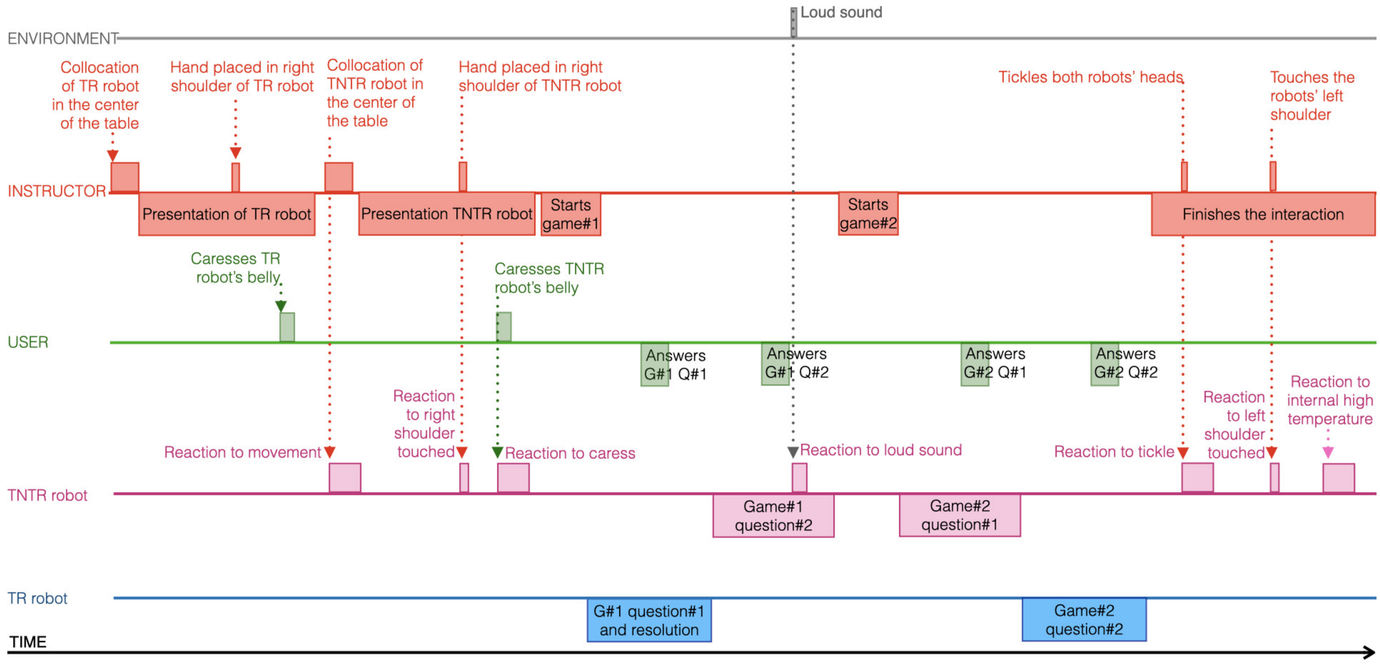


Fig. 4. Timeline of the interaction in the videos

tionnaire [20]. This questionnaire measures people judgments of different social attributes of robots through 18 adjectives, as follows: dangerous, awkward, aggressive, feeling, strange, knowledgeable, reliable, happy, compassionate, awful, competent, social, responsive, scary, capable, emotional, interactive and organic (non-mechanical). The adjectives are used to compute a scale that reveals the factor's warmth, competence, and discomfort. These three factors allow for independence from the appearance of the robot because the RoSAS scale was made based on the concept of 'robot' and not on examples of concrete robots.

For each robot, the participant is presented with a picture of the robot extracted from the video, and they are asked to rate on a scale of 1 to 7 about how well the 18 adjectives describe each of the robots. On this scale, 1 corresponds to 'not at all' and 7 corresponds to 'very much so'. The survey was made in Spanish, so these adjectives were translated from the original English survey. These adjectives, although shown randomised to avoid bias, are then used to compute the three factors that the RoSAS scale assesses.

To conclude the survey, the participation is thanked and a contact e-mail is provided. They are also given the option to leave a comment.

E. Hypotheses

In this experiment, we seek to validate the following hypotheses:

- **H0**: The addition of reactions to non-task-related events will affect how the robot is perceived.
- **H1**: A robot that is capable of reacting to non-task-related events will be perceived as being warmer than a robot that does not react to those events.
- **H2**: A robot that is capable of reacting to non-task-related events will be perceived as being more competent than a robot that does not react to those events.
- **H3**: A robot that is not capable of reacting to non-task-related events will be perceived as being more uncomfortable than the robot that does react to those events.

F. Participants

A total of 182 participants were recruited through various social media, in particular WhatsApp, Instagram and Twitter. They participated voluntarily, without any form of compensation.

TABLE II
DISTRIBUTION OF THE PARTICIPANTS IN THE EXPERIMENT

Category	Sub-category	Number of participants
Gender	Female	82
	Male	61
Age	< 20	5
	20 to 29	63
	30 to 39	10
	40 to 49	15
	50 to 59	41
	$>= 60$	9
Education	Primary Education	4
	High School	15
	Vocational training	19
	Bachelor's degree	71
	Master's degree	33
	PhD	1

Before performing the statistical analysis, we analysed the responses obtained to identify invalid responses. First, the data from seven participants were deleted because they answered that they had not watched the video until the end. We considered that they would probably not have watched the two conditions completely and their answer could alter the reliability of the results. Next, 24 duplicated answers were identified and deleted. Finally, we used the SPSS software to conduct our statistical analysis. This program includes a feature that identifies outliers using Tukey's method [31]. Through this method, eight extra answers were identified as anomalous. We double-checked them manually and observed that they were responses where the participant had rated every attribute in the RoSAS scale exactly the same. This led us to discard them. In total, data from 143 participants were used for the statistical analysis.

Of these 143 participants, 57.34% identified themselves as female and 42.66% as male. The ages of the participants ranged from 18 to 77 years old. In total, 90.91% live with a partner and 9.09% live alone. Regarding their education, 50% of them have a Bachelor's degree.

V. Results

The ratings for each of the factors of the RoSAS questionnaire for both conditions have been compared using the Wilcoxon Signed-rank Test [32]. This test was chosen because several of the assumptions of the Paired Samples T-test [33], such as assuming the normality of the samples, could not be affirmed because the data collected did not follow a normal distribution. The Wilcoxon Test assumes that the dependent variables could be measured at the ordinal level (in this case, they have been rated on a range from 1 to 7) and that the same subjects are present in both groups, this condition is fulfilled by each participant having observed both conditions (the TNTR and TR robots). Finally, the distribution of both groups, although not normal, does have the same shape.

The results indicate that the differences in the ratings for warmth between the TNTR and TR robot are statistically significant (Z=-6.171, $p<0.001$), as were the ratings for competence between the two conditions (Z=-4.195, $p<0.001$). However, the difference between the ratings of discomfort for the two robots was not statistically significant (Z=-0.016, p=0.987). These results are shown in Table III. When also considering demographic factors, such as age or gender, we did not observe significative differences for any of the three factors. The population means for factors warmth and competence for both conditions are shown in Table IV, and are graphically presented in Figure 5.

TABLE III
Results of the Wilcoxon Signed Rank Test for the three factors

RoSAS factor	Z	p-valor
Warmth_{TNTR} Warmth_{TR}	-6.171	<.001
Competence_{TNTR} Competence_{TR}	-4.195	<.001
Discomfort_{TNTR} Discomfort_{TR}	-.016	.987

Finally, if we check the questions that asked the participants about their preference between both robots, then we can see that a majority of them selected the TNTR robot over the TR one for every question. In total, 65.03% of participants reported a preference to play with the TNTR robot, 64.33% would rather have the TNTR robot at home, 67.13% would rather be accompanied by the TNTR robot, and 63.64% of participants felt more comfortable with the TNTR robot.

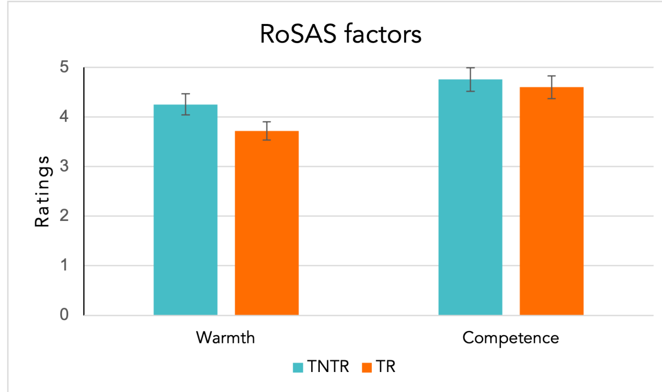


Fig. 5. Ratings for the warmth and competence factors (both showed significant differences between conditions). Bars represent the mean rating, while the whiskers represent the confidence interval at 95%

TABLE IV
Statistics for the three factors

RoSAS factor	Median	Deviation	Percentil 50 (Median)
Warmth_{TNTR}	4.25	1.50	4.50
Warmth_{TR}	3.72	1.49	3.50
Competence_{TNTR}	4.76	1.31	4.83
Competence_{TR}	4.60	1.32	4.67
Discomfort_{TNTR}	1.90	0.99	1.67
Discomfort_{TR}	1.88	0.89	1.67

VI. Discussion

After analysing the results, the significant differences with regard to the warmth and competence factors highlight the improvement of how the participants perceived the robot under the TNTR condition. Two of the three factors have shown significant differences. Therefore, it can be said that the main hypothesis, H0, is validated. The fact that there is a significant difference in the ratings for warmth and competence validates hypotheses H1 and H2. With the addition of the non-task-related reactive behaviours, participants did perceive the robot to be warmer, which validates H1. Meanwhile, in the case of H2, the observed increase in the competence may happen because the robot under the TNTR condition is able to detect more events in the environment and react to them, which could lead to a perception of a higher intelligence and make the robot appear to be more competent. These results are in line with those reported by Birnbaum et al. [25], where the robot under the reactive condition was found to be more competent than the robot under the non-reactive condition. These results suggest that the implementation of non-task-related reactive gestures that increase the liveliness of the robot has helped to improve the participants' perceptions.

Although hypothesis H3 is not validated because there are no significant differences for the ratings of the discomfort factor, a positive conclusion can be extracted from this. A non-significant difference for this factor might indicate that by introducing the implemented gestures, negative connoted features have not been perceived with a higher degree by the participants when the robot is reactive to non-task-related events. This factor is very important in social robots because being close with people might improve the human-robot interaction. In addition, it is important to remark that the discomfort rating is low for both conditions. A possible reason for this is the fact that the robot's appearance was designed to be friendly and likeable. The robot was also able to conduct an interaction appropriately under both conditions, which could also have had an effect on the discomfort ratings. Further tests would be necessary to prove these hypotheses. Finally, it was observed that a majority of participants would rather play with (65%), have at home (64.33%), be accompanied by (67.13%), and be more comfortable with (63.64%) the robot under the TNTR condition. This provides further proof that the addition of the non-task-related reactions improved how the participants perceived our robot.

The fact that the TNTR robot is perceived as warmer suggests that the reactive behaviours that we implemented make the participant perceive it in a more friendly and interactive way. This might mean that the interaction between the TNTR robot and humans will be more fluid and dynamic because it is more human-like. In addition, by perceiving the robots as interactive beings, their acceptance might be easier than if the participant perceives them to be cold and distant. However, further tests would be necessary to evaluate if these hypotheses are true. Furthermore, the fact that the discomfort factor has not increased leads us to think that the robot's reactions have simply increased the positive aspects of the robot's appearance, and have no negative effects on the observer's perception.

In summary, the results indicate that when our robots react to non-task-related events, they are perceived more positively than when they do not. These results are in line with the work of Rifiniski et al. [23], who demonstrated that intelligence, animacy and likeability of the robot are higher when it is reactive, and with the study of Birnbaum et al. [25], where the robot Travis was perceived as more reactive, social and competent in the reactive condition. Furthermore, in the study of Hoffman et al. [22], it was shown that robots are rated as more attractive when they react. Therefore, for the design of new robots, it would be interesting to include this type of behaviour or to expand the range of reactions in current robots because they may help the robot to be rated more positively by the participants.

A. Limitations

Several limitations were observed during the evaluation of the behaviours that we implemented. First, we were forced to proceed with a video-based evaluation because of the COVID-19 pandemic. Compared to usual techniques where the evaluated user interacts directly with the robots, it is true that with screen-based methods the accuracy of the results may decrease because it is not possible to know for sure if the participant has watched the video completely or if they were paying attention. Furthermore, the participant cannot be fully aware of the size of the robot or the real sounds [27]. It is also difficult to reach all population groups because not everyone uses social media.

Language was another limitation. The evaluation was made in Spain, where the main language is Spanish. Therefore, we had to translate the adjectives of the RoSAS questionnaire because some of our participants do not speak English. Furthermore, some of the adjectives do not have a literal translation, so its meaning may have been slightly altered. The questionnaire is validated in English, and a validation in Spanish would be needed to ensure the accuracy of the results. The robots that we used for the evaluation were a third limitation. While we prepared two different videos, where the colour of the TNTR and TR robots was switched, the head and arms of each robot did not change. Thus, the TNTR robot always had a fluffy head, while the TR one had a solid head. This could have affected the perception of the user because the TNTR robot might have been seen as softer, or even more huggable. Further tests would be required to evaluate if the addition of a fluffy exterior to the robot's head and arms has an effect on how it is perceived by the users. Finally, the positioning of the robots was kept unchanged, with the TR robot placed to the right-hand side and the TNTR robot placed to the left-hand side. This might create an issue because of the way in which humans process visual stimuli. In the area of human vision, there is a concept known as visual field asymmetries, which describes differences in the processing of stimuli depending on which part of the visual field it appears [34]. In particular, the left/right visual field asymmetry may have been an issue. For example, the results reported by Michael and Ojéda [35] suggest the existence of left/right hemispheric asymmetry in selective attention (i.e., trying to pay attention to specific information while ignoring non-relevant stimuli). However, they reported that the asymmetry was more evident with high stimuli similarity, and thus task demands. Consequently, it would be necessary to test if the placement of the robots has had an effect on how the participants perceived each of the robots.

Finally, when designing the verbal part of the reactions, we drew inspiration from sentences that are commonly used by Spanish speakers when reacting to the stimuli considered. Therefore, some of the sentences can be interpreted as emotional, and thus could have an effect on how the users perceive the TNTR robot.

More evidence would be needed to validate the impact of these limitations because we cannot say if they affected the results of the evaluation.

VII. Conclusion

This paper offers some findings about using the human perception of reactive robots as a way to improve human-robot interaction. Our results could bring some light on the design of social robot's behaviours and how stimuli reaction can make robots more human-like.

The principal conclusion of the work is that we can achive more lively robots by mamking them to react to stimuli not related to the activity the robot is actively engaged, even if the robot was already able to react appropriately to events tied to the task at hand. It should be noted that these behaviours were designed based on typical human reactions and the way in which we respond to stimuli. We tried to imitate the way that we receive the stimuli by using the same interface or an equivalent. For example, capacitive sensors simulate the receptors that we have in our skin, and both the eyes and the different parts of the body are used to give responses. While the robots have certain limitations, they have been overcome with the use of additional elements (e.g., the heart colours).

The software has been implemented in such a way that it is platform-independent, regardless of the hardware elements they have. It is also modular, which would make it easy to expand if necessary. Currently, the system has been integrated in three robotic platforms. Future lines of work include the addition of new stimuli that the robot can react to and also to test the proposed system in more robots. It would also be interesting to evaluate two other factors. First, it would be interesting to evaluate if a bad design of these reactions could have a negative effect over how participants perceived the robot. Having the robot react in unnatural ways might not improve, or even negatively affect, the perception of a user when compared with a robot that shows a complete lack of reactiveness. The second interesting evaluation is to test if there is a need for the robot's non-task-related responses to be designed in a human-like fashion or if more robotic reactions would improve the user's perception of the robot, and how they would perform when compared with the human-like behaviours.

After an evaluation with participants, it can be concluded that the impact of the developed behaviours improves the warmth and competence of the robot, while keeping the user's perception of discomfort at low levels.

REFERENCES

[1] M. Blow, K. Dautenhahn, A. Appleby, C. Nehaniv, and D. Lee, "The art of designing robot faces: dimensions for human-robot interaction," 01 2006, pp. 331–332.

[2] H. Robinson, B. Macdonald, N. Kerse, and E. Broadbent, "The psychosocial effects of a companion robot: A randomized controlled trial," *Journal of the American Medical Directors Association*, vol. 14, 03 2013.

[3] J. Gorostiza, R. Barber, A. Khamis, M. Pacheco, R. Rivas, A. Corrales-Paredes, E. Delgado, and M. Salichs, "Multimodal human-robot interaction framework for a personal robot," 10 2006, pp. 39 – 44.

[4] K. E. Powers, A. L. Worsham, J. B. Freeman, T. Wheatley, and T. F. Heatherton, "Social connection modulates perceptions of animacy," *Psychological science*, vol. 25, no. 10, pp. 1943–1948, 2014.

[5] B. Reeves and C. Nass, "The media equation: How people treat computers, television, and new media like real people," *Cambridge, UK*, vol. 10, p. 236605, 1996.

[6] C. Nass, J. Steuer, and E. R. Tauber, "Computers are social actors," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1994, pp. 72–78.

[7] G. Csibra, G. Gergely, S. Bíró, O. Koós, and M. Brockbank, "Goal attribution without agency cues: the perception of "pure reason" in infancy," *Cognition*, vol. 72, no. 3, pp. 237–267, 1999.

[8] E. Fromm, *The Heart of Man: Its Genius for Good and Evil*. New York, NY: Harper & Row, 1964.

[9] P. H. Kahn, "Developmental psychology and the biophilia hypothesis: Children's affiliation with nature," *Developmental Review*, vol. 17, no. 1, pp. 1 – 61, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S027322979690430X

[10] R. Ulrich and O. Lunden, "Effects of nature and abstract pictures on patients recovering from open heart surgery," in *International Congress of Behavioral Medicine*, jun 1990.

[11] Y. Joye and A. De Block, "'nature and i are two': A critical examination of the biophilia hypothesis," *Environmental Values*, pp. 189–215, 2011.

[12] J. Stacey and L. Suchman, "Animation and automation – the liveliness and labours of bodies and machines," *Body & Society - BODY SOC*, vol. 18, pp. 1–46, 03 2012.

[13] I. Jara-Rivera, A. Castro-González, J. Castillo, F. Alonso-Martín, M. Malfaz, and M. Salichs, "Creating lively behaviors in social robots," 10 2017, pp. 283–292.

[14] S. Rossi, A. Rossi, and K. Dautenhahn, "The secret life of robots: Perspectives and challenges for robot's behaviours during non-interactive tasks," *International Journal of Social Robotics*, vol. 12, 12 2020.

[15] B. Duffy, "Anthropomorphism and the social robot," *Robotics and Autonomous Systems*, vol. 42, pp. 177–190, 03 2003.

[16] A. 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, 03 2016.

[17] A. Sowmya, "Real-time reactive model for mobile robot architecture," K. W. Bowyer, Ed., pp. 713–721. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=992619

[18] T. Sun, Z. Dai, and P. Manoonpong, "Distributed-force-feedback-based reflex with online learning for adaptive quadruped motor control," *Neural Networks*, vol. 142, pp. 410–427, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608021002331

[19] G. Hoffman, O. Zuckerman, G. Hirschberger, M. Luria, and T. Shani-Sherman, "Design and evaluation of a peripheral robotic conversation companion," vol. 2015, 03 2015.

[20] C. Carpinella, A. Wyman, M. Perez, and S. Stroessner, "The robotic social attributes scale (rosas): Development and validation," 03 2017, pp. 254–262.

[21] D. Cameron, S. Fernando, E. Collins, A. Millings, R. Moore, A. Sharkey, and A. Prescott, "Impact of robot responsiveness and adult involvement on children's social behaviours in human-robot interaction," 06 2016.

[22] G. Hoffman, G. Birnbaum, H. Reis, K. Vanunu, and O. Sass, "Robot responsiveness to human disclosure affects social impression and appeal," 03 2014.

[23] D. Rifinski, H. Erel, A. Feiner, G. Hoffman, and O. Zuckerman, "Human-human-robot interaction: robotic object's responsive gestures improve interpersonal evaluation in human interaction," *Human-Computer Interaction*, vol. 36, pp. 1–27, 02 2020.

[24] G. Hoffman, S. Bauman, and K. Vanunu, "Robotic experience companionship in music listening and video watching," *Personal and Ubiquitous Computing*, vol. 20, pp. 1–13, 02 2016.

[25] G. Birnbaum, M. Mizrahi, G. Hoffman, H. Reis, E. Finkel, and O. Sass, "Machines as a source of consolation: Robot responsiveness increases approach behavior and desire for companionship," 03 2016.

[26] M. Salichs, A. Castro-González, E. Salichs, E. Fernández-Rodicio, M. Maroto Gómez, J. J. Gamboa, S. Marques, J. Castillo, F. Alonso-Martín, and M. Malfaz, "Mini: A new social robot for the elderly," *International Journal of Social Robotics*, vol. 12, 12 2020.

[27] G. Venture and D. Kulic, "Robot expressive motions: A survey of generation and evaluation methods," *ACM Transactions on Human-Robot Interaction*, vol. 8, pp. 1–17, 11 2019.

[28] G. Charness, U. Gneezy, and M. A. Kuhn, "Experimental methods: Between-subject and within-subject design," *Journal of economic behavior & organization*, vol. 81, no. 1, pp. 1–8, 2012.

[29] R. S. Jhangiani, I.-C. A. Chiang, C. Cuttler, D. C. Leighton *et al.*, *Research methods in psychology*. Kwantlen Polytechnic University, 2019.

[30] C. K. Hsee, G. F. Loewenstein, S. Blount, and M. H. Bazerman, "Preference reversals between joint and separate evaluations of options: a review and theoretical analysis." *Psychological bulletin*, vol. 125, no. 5, p. 576, 1999.

[31] J. W. Tukey, *Exploratory Data Analysis*. Addison-Wesley Pub. Co.

[32] (2021, 05) Wilcoxon signed-rank test using spss statistics. [Online]. Available: https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php

[33] (2021, 05) Dependent t-test using spss statistics. [Online]. Available: https://statistics.laerd.com/spss-tutorials/dependent-t-test-using-spss-statistics.php

[34] E. S. Gorbunova and M. V. Falikman, "Visual search for letters in the right versus left visual hemifields," *Advances in Cognitive Psychology*, vol. 15, no. 2, p. 75, 2019.

[35] G. A. Michael and N. Ojéda, "Visual field asymmetries in selective attention: Evidence from a modified search paradigm," *Neuroscience Letters*, vol. 388, no. 2, pp. 65–70, 2005.