RoSI: A Model for Predicting Robot Social Influence

HADAS EREL*, Media Innovation Lab, Reichman University, Israel
MARYNEL VÁZQUEZ, Department of Computer Science, Yale University, USA
SARAH SEBO, Department of Computer Science, University of Chicago, USA
NICOLE SALOMONS, I-X and the Department of Computing, Imperial College London, United Kingdom
SARAH GILLET, Division of Robotics, Perception and Learning, KTH Royal Institute of Technology, Sweden
BRIAN SCASSELLATI, Department of Computer Science, Yale University, USA

A wide range of studies in Human-Robot Interaction (HRI) has shown that robots can influence the social behavior of humans. This phenomenon is commonly explained by the Media Equation. Fundamental to this theory is the idea that when faced with technology (like robots), people perceive it as a social agent with thoughts and intentions similar to those of humans. This perception guides the interaction with the technology and its predicted impact. However, HRI studies have also reported examples in which the Media Equation has been violated, that is when people treat the influence of robots differently from the influence of humans. To address this gap, we propose a model of Robot Social Influence (RoSI) with two contributing factors. The first factor is a robot's violation of a person's expectations, whether the robot exceeds expectations or fails to meet expectations. The second factor is a person's social belonging with the robot, whether the person belongs to the same group as the robot or a different group. These factors are primary predictors of robots' social influence and commonly mediate the influence of other factors. We review HRI literature and show how RoSI can explain robots' social influence in concrete HRI scenarios.

CCS Concepts: • **Human-centered computing → HCI theory, concepts and models**.

Additional Key Words and Phrases: Human-Robot Interaction, Social Influence, Expectation, Belonging

ACM Reference Format:
Hadas Erel, Marynel Vázquez, Sarah Sebo, Nicole Salomons, Sarah Gillet, and Brian Scassellati. 2024. RoSI: A Model for Predicting Robot Social Influence. 1, 1 (May 2024), 22 pages. https://doi.org/10.1145/3641515

1 INTRODUCTION

A wide range of experimental work in Human-Robot Interaction (HRI) has shown that robots can influence the social behavior of humans. For example, prior work suggests that the mere presence of robots can facilitate honest human behavior [40], that robot nonverbal behavior can induce changes in human nonverbal behavior as well [49], and that robots can motivate people to engage in activities like exercising [30, 73]. Further, robot cheating behavior has been

*All authors contributed equally to this research.

Authors' addresses: Hadas Erel, Media Innovation Lab, Reichman University, Herzliya, Israel, hadas.erel@milab.idc.ac.il; Marynel Vázquez, Department of Computer Science, Yale University, New Haven, USA, marynel.vazquez@yale.edu; Sarah Sebo, Department of Computer Science, University of Chicago, USA, sarahsebo@uchicago.edu; Nicole Salomons, I-X and the Department of Computing, Imperial College London, London, United Kingdom, n.salomons@imperial.ac.uk; Sarah Gillet, Division of Robotics, Perception and Learning, KTH Royal Institute of Technology, Sweden, sgillet@kth.se; Brian Scassellati, Department of Computer Science, Yale University, USA, scaz@cs.yale.edu.

found to influence human engagement in human-robot interactions [88] and robot emotional expressions may influence human willingness to collaborate with these machines [104]. In group settings, robots have further been shown to influence not only human-robot interactions, but also human-human interaction dynamics [20, 27, 35, 75, 79, 91].

For 25 years, our understanding of robot social influence has been dominated by Reeves and Nass's Media Equation [74]. The fundamental premise of the Media Equation is that when predicting how people will act when interacting with a novel technology, one needs only to look at how people normally engage with each other. If wearing identical blue armbands makes two people act more like a team, then we predict a similar outcome when a person and a robot share the same armbands. This idea has had a foundational influence on both the fields of Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI). We explicitly design human-computer and human-robot interactions by following examples provided by human-human interactions with the strong assumption that these will continue to hold true (e.g., [5]).

There are times, though, when the Media Equation does not offer a prediction because there is no equivalent in human-human behavior. We can make no prediction about how a person will behave when asked to turn a robot off and then on again as this is not something that could be done to a person. Perhaps more importantly, there is a growing collection of results that show cases where the prediction made by the Media Equation falls flat [33]. People often feel comfortable making mistakes in front of robot tutors when they would not let these errors be seen by human tutors [54]. We accept that robots can take on roles that we do not consider appropriate for people [52]. We also feel pressure to conform to groups of humans, even when they give obviously incorrect answers to simple questions, but not to robots [13].[1]

Many researchers have tried to address the limitations of the Media Equation; however, these alternate models focus on predicting how people will perceive technology and do not predict the type and magnitude of the social influence exerted by the technology. For instance, several alternative models focus on computers [85] or voice/virtual agents [67]. They suggest that people use different communication strategies when interacting with technology, use different cognitive processes, and are less engaged. Several models define interaction with technology as a separate independent type of interaction that involves novel social scripts, triggering different types of control and confidence [33, 67]. Specifically, in the context of robots, a few models were suggested to address the complexity of how people perceive robots. For example, Zlotowski et al. (2018) suggested that human perceptions of a robot depend on a dual evaluation process that involves both an implicit automatic perception and an explicit controlled evaluation of the robot [105]. Ruijten et al. (2019) suggested an alternative that involves the ordering of human-like characteristics on a range of perceived human-likeness [78]. Another recent explanation for robot perception was suggested by Clark and Fisher (2023). They argue that people do not perceive robots as social agents themselves but that they construe robots as interactive depictions [18]. The authors suggest that robots are perceived from three different perspectives: (1) Base Scene - the robot's physical aspects (appearance, material, design); (2) Depiction Proper - the features one associates with the base scene, such as what social agent the robot represents; (3) Scene depicted - based on the previous two perspectives, the robot is perceived as a character, which allows for imagining probable scenes involving the robot. The authors suggested that people can effortlessly take any of these perspectives and switch between them. Other theories have emphasized the human-likeness of the robot's appearance, behavior, social skills, and level of personalization as factors that determine how people perceive robots [26, 43]. While these models can account for the limitations of the Media Equation in explaining how people perceive robots, they do not demonstrate a difference in the magnitude of the

[1]Though this conformity effect may be seen in more subtle ways [82].

impact on people's social behavior. If people do not perceive robots similarly to people it is important to identify the similarities and differences of such robotic social influence.

In this paper, we propose a model for predicting the magnitude of a *robot's social influence* through two independent factors: *violation of expectations* and *belonging*. Our model focuses on what we consider to be the two most important and critical factors, at times valuing simplicity over complete accounting of all possible factors. We believe that these two factors capture the unique social context that a robot affords as something that is at times an inanimate object (and not subject to the classical Media Equation) and, at times, a social agent (i.e., the interaction with it is characterized by human-like social effects, [41]), creating a specific set of expectations around a particular interaction. Furthermore, the robot's status as being in some ways similar to ourselves (as an in-group member) and at times very different from ourselves (as an out-group member who does not belong) shapes the influence that these artifacts have over us.

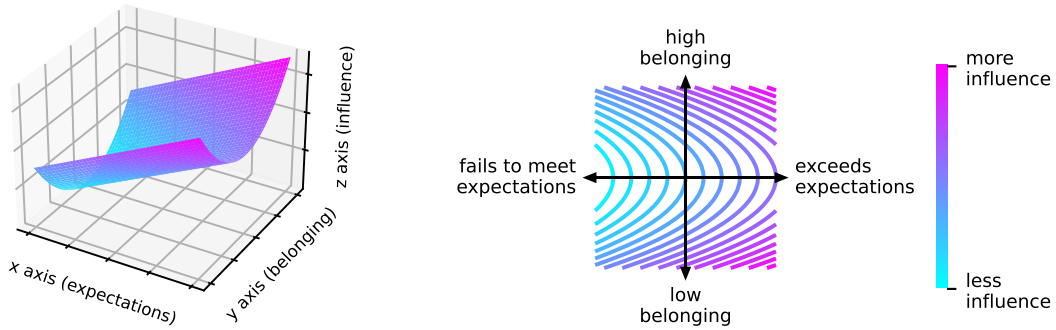
In the following sections, we first explain and ground our three factors: the dependent and predicted factor of *social influence* and the independent factors of *violation of expectation* and *belonging*. Then, we provide the first step of validating our model, which we call RoSI (Robot Social Influence), by applying the model to five previously published user studies and predicting their results. We visualize these predictions in use-case diagrams. Through those example predictions, we provide insight into how RoSI should be applied to HRI experimental designs. The goal of this work is to provide a theoretical model that can explain the outcome of an experiment by predicting differences in the robot's social influence.

2 A MODEL OF ROBOT SOCIAL INFLUENCE (RoSI)

Figure 1 describes our theoretical model of Robot Social Influence (RoSI), which predicts the magnitude of a robot's social influence on a person based on two factors: (1) the degree to which the robot exceeds or fails to meet the person's expectations and (2) the degree to which the person considers themselves and the robot as belonging to the same group. In particular, the model predicts a larger magnitude of social influence when the robot exceeds expectations and a lower magnitude when it fails to meet them. It also predicts that both high and low belonging will involve high social influence, while a more neutral sense of belonging will involve a low influence.

We chose violations of expectations and belonging specifically as independent variables of our model because they are primary components and powerful predictors of robot social influence across a broad range of human-robot interactions. In some cases, these two factors are predicted to involve different influence patterns for humans and robots (e.g., our expectations of robots can differ from our expectations of people). Previous studies have indicated other factors that can also shape the social influence of a robot, including robot anthropomorphism [69, 103], robot status or power [81], robot competence [38], social pressure [13], and perceived robot agency [56, 88]. These other factors are commonly encapsulated in our chosen factors of expectations (e.g., power, competency, agency) and group belonging (e.g., robot status, anthropomorphism).

The model makes several key assumptions: (1) social influence is presented from the perspective of a person who takes part in an interaction with a robot; (2) the model predicts the magnitude, but not the valance or type of the social influence (which are highly context-dependent); (3) the model predicts the change in the magnitude of social influence across the interaction with the robot. While at the beginning of an interaction, people have initial expectations and sense of belonging, the robot's behavior throughout the interaction can alter those, which would result in a different magnitude of social influence.



(a) 3D visualization (b) 2D projection

Fig. 1. RoSI, our social influence model, predicts the amount of influence a robot has on a person (z axis) based on: the degree to which the robot violates the person's expectations (x axis), and the degree to which the person considers themselves and the robot to belong to the same group (y axis). The 3D plot (a) shows the relationship between these axes. The 2D plot (b) shows a 2D projection of the social influence surface from (a). Both diagrams are best viewed in color.

2.1 Predicted Factor: Robot Social Influence

Prior work in HRI has demonstrated that robots can change how people behave, how people perceive the world around them, and how people think and feel. In other words, robots can exert social influence on people. We define social influence as "a change in one's beliefs, behavior, or attitudes due to external pressure that may be real or imagined" [37]. Social influence can take many forms: a person may conform their choice to match that of a robot [82, 83, 98], a person may comply with the request of a robot (even if it does not make sense) [6, 76], or a person may follow a social norm introduced or reinforced by a robot [68, 93].

2.1.1 Theoretical Background & Justification. Rooted in the human sensitivity to social information in their environment [61], people perceive robots' actions as social and, as a result, are often influenced by them. It is argued that humans are social organisms [94], with brain structures that specifically support awareness of others and communication skills that are fundamental to social coordination [51, 61, 74]. The inherent propensity to perceive the world through a social lens [101] is believed to lead to a strong tendency to anthropomorphize objects and to automatically associate autonomous actions with social intent [2, 25, 29, 51, 90]. This social interpretation given to interactions with autonomous objects is the basis for the various indications for social influence in human-robot interactions.

At the same time, robots do not necessarily have the same social influence as humans do (e.g. [13]). Different aspects of a robot's design may shape the type and magnitude of its social influence. The robot's appearance, capabilities, communication modalities, and role may all contribute to its social influence in a given interaction. We suggest that it is not enough to state that interactions with robots are interpreted as social experiences. Therefore, we present this model of robot social influence (RoSI), that can be used to predict the unique social influence of a robot on the people with whom the robot interacts.

2.2 Model Independent Factor #1: Violation of a Person's Expectations of the Robot

The first factor in RoSI is how much and in which direction a robot violates a person's *social* expectations. Social expectations are defined as cognitions concerning the behavior anticipated of others when interacting with them [4, 16]. They guide people's behavior by reducing uncertainty and directing their comprehension [11, 14, 16, 84]. In the context of human-robot interaction, expectations become even more important due to the inherent uncertainty associated with an autonomous machine [90]. A robot could meet the person's expectations, exceed their expectations, or fail to meet the person's expectations. A robot can also violate the person's expectations in a way that surprises the person but does not exceed or fail to meet their expectations. We treat these cases as similar to the robot meeting the person's expectations.

For example, consider a person's expectations of a food-delivery robot. One way the robot might exceed expectations is by displaying empathy in addition to delivering food, for example, noticing that a person looks sad and saying "you look sad, is everything ok?" A food-delivery robot could fail to meet expectations if it is unable to speak its "here's your delivery!" phrase as the user has come to expect due to a software failure. A robot might surprise a person by using a voice that the person did not expect but neither exceeds expectations or fails to meet them.

Robots that exceed a person's expectations typically present complex social behaviors like vulnerability [92], favoring one human over the other [45], and touching the human [19]. Exceeding expectations does not necessarily involve positive robotic behavior as some negative behaviors like cheating and excluding others [28, 56] may be perceived as exceeding the robot's anticipated behavior. Robots that fail to meet a person's expectations often disappoint them since they exhibit lower capabilities than expected [47]. Robots fail to meet expectations either due to humans' unrealistic initial expectations or due to mistakes and errors related to the robot's function (e.g., navigation errors [80], perceptual and processing judgment errors [47], memory-related errors [102]).

While some robot behaviors may cause expectation violations that always exceed a person's expectations or always fail to meet them, some robot behaviors may cause distinct effects based on individual differences. For example, consider the possibility that a food-delivery robot receives a software update that has the robot express happiness after a successful delivery by spinning around in a circle. This additional expression of emotion could exceed a person's expectations of the robot. However, another person may perceive the circle-spinning as a robot malfunction, resulting in a failure for the robot to meet the person's expectations. It is also possible that the circle-spinning of the robot may not change a person's perceptions of the robot, leaving that person's expectations of the robot unchanged (leaving their expectations at "meets expectations"). As this example illustrates, the violation of expectations can be subjective.

2.2.1 Theoretical Background & Justification. Violation of the expectations from the robot was chosen as one of the model's factors due to the centrality of expectations in interpersonal interactions [14, 72, 90] and specifically how they shape human behavior in HRI [48]. One of the primary concerns associated with interpersonal interactions is the uncertainty about others' thoughts, attitudes, and behavior. It is suggested that in any given context, people unconsciously develop expectations that assist in predicting different aspects of the interaction [4]. Such predictions have a persistent influence on the social sensitivity in the interaction [84, 90].

Expectation may be even more central in human-robot interactions [72] because human norms cannot be fully applied to interactions with robots [58], and people commonly have various and changing expectations of robots [77]. At the same time, robots' autonomous behavior typically positions them as independent social actors, triggering a set of expectations people apply to social contexts. Accordingly, expectations from robots are formed by various factors ranging from relevant human norms in the social context to experience with real and fictional robots. Technical affinity

and the introduction of a robot in a specific interaction were also shown to impact the expectations from the robot [36, 44, 70]. This complexity increases the need to understand and anticipate robots' behavior during interactions, emphasizing the importance of deriving accurate expectations. When expectations are violated, one is required to reassess the situation which leads to higher social sensitivity. The direction of the violation (exceeding or failing to meet expectations) and its magnitude determine the intensity of the social impact [14, 15]. Various studies have already mapped people's expectations of different robots (e.g. [77]) and suggested methods for manipulating, structuring and modifying expectations. These commonly involve a short explanation about the robot given before the interaction [36, 70].

2.2.2 Prediction of Social Influence. Expectations are constructed in RoSI as a relative factor predicting the intensity of the robot's social influence. The model treats each person's expectations violation as relative to their own initial expectations. Thus, the starting point for a person's expectations of the robot is at the center of the horizontal axis of Fig. 1, where the robot meets the expectations.

When a robot meets the person's expectations, the model predicts no change in the robot's social influence. No change in the robot's influence is also predicted when the violation of expectations is surprising since it neither exceeds the person's expectation nor fails to meet the person's expectations. While a surprise at a robot's actions may influence the human-robot interaction in some way, we only expect a significant change in the robot's social influence when the person's expectations of the robot are either exceeded, or the robot fails to meet the person's expectations..

Robotic experiences that fail to meet expectations are predicted to decrease social influence (due to disappointment and decrease in trust [99]). On the other hand, robotic experiences that exceed participants' expectations are predicted to increase social influence (due to the increase in the robot's perceived capabilities). While the magnitude of the social influence is predicted to increase, it is not necessarily predicted to be positive. For instance, a robotic experience that exceeds expectations may lead to increased conformity and obedience.

The relative nature of this factor also suggests that the robot's social influence will become more typical over repeating interactions since the expectations from the robot will be updated. If the robot presents consistent behaviors across several interactions, the expectations from it will be adjusted accordingly. This, in turn, would form a relevant social context with a typical social influence.

As an illustrative example, consider a robot that provides product recommendations. Accepting the robot's recommendation for a product to purchase is predicted to depend on the robot's perceived group membership and the initial expectations from the robot (whether the robot meets them). If the robot fails to meet expectations (e.g., unable to provide information about the product or if there are long delays in its responses to questions), the chances that its recommendation will be followed drop. On the other hand, if the robot provides a sense of caring by asking for the person's name and preferences, it is likely to be perceived as exceeding expectations, and its recommendations are more likely to be accepted by the user.

2.3 Model Independent Factor #2: A Person's Belonging to the Robot's Group

The second input to RoSI is the person's perception of their group membership relative to the robot: whether the person and the robot belong to the same group or different groups. Let us consider a card game where two human-robot teams play competitively against each other: participant A and robot Emys play against participant B and robot Glin, as explored by Correia et al. [22]. Participant A would likely perceive a high degree of belonging to robot Emys's group (their partner) and a low degree of belonging to robot Glin's group (their competitor). We view high and low belonging

to the robot's group as a similar concept to intergroup membership (in-group/out-group) [95]. High belonging to the robot's group is analogous to the person viewing the robot as an in-group member. Low belonging to the group is analogous to the person viewing the robot as an out-group member. It is also possible for a person's belonging to the robot's group to exist somewhere in between low belonging and high belonging (e.g., at a neutral midpoint).

People's perceptions of their and the robot's group membership can be powerfully shaped by both the robot's behavior and environmental factors. People show greater signs of shared group membership (e.g., favor, trust, liking) with robots that express vulnerability [60, 92], use team-related verbal expressions [21, 81], display empathy [53], and use social touch [50, 87, 100]. Robots do not even need to have a humanlike appearance to engender themselves to greater group membership and belonging with people. Using expressive gestures, non-humanoid robots have effectively communicated responsiveness [12] and security [59] in one-on-one interactions with people. In addition to robot behavior, environmental factors, such as pre-assigned interaction roles [21, 32, 91], can also influence people's views of robots' group membership. For example, when robots are a part of the same team as a person, those robots are seen as in-group members, however, when robots are a part of a competing team to a person, those robots are seen as out-group members [32].

2.3.1 Theoretical Background & Justification. The focus on belonging as one of the two factors determining the robot's social influence is grounded in its profound impact on human behavior in interpersonal interactions [10, 17]. The dramatic effect of belonging is attributed to a fundamental human drive to form meaningful relationships with others that involve positive and pleasant interactions [1, 9, 10]. In order to satisfy the need to belong, people must establish meaningful interpersonal connections [63] commonly supported by sharing group membership [10, 42]. Belonging to a group leads to the rapid formation of strong group bonds, loyalty, and group identification ties [42, 57, 86, 86]. Sharing group membership encourages behaviors that enhance the chances of being included, such as showing favoritism to in-group members [7, 42, 57, 86] and defending the integrity of the intergroup social bonds [10]. Belonging to a group also influences the perceptions of out-groups [10, 62] leading to negative attitudes and rejection of those with different group membership [62]. This variety of group membership effects contributes to the centrality of belonging in shaping interactions.

While group membership effects are also observed in HRI, they are not always similar to those observed in human interactions and present more complex patterns of social influence that underscore their importance to the model. Robots are not naturally perceived as in-group members and can often be considered a part of a potentially competing out-group [89]. People are more likely to consider themselves as members of a group with people who share more in common with them [24]. This is supported by evidence that people demonstrate greater in-group favoritism to other humans as opposed to robots [32], and to more human-like robots as opposed to more machine-like robots [31]. In some cases, interactions with robots fail to show any belonging effects that are typically observed in human groups [13]. However, robots' behavior and various environmental factors may form a strong sense of mutual human-robot group membership. In such contexts, people have shown a preference for in-group robots over out-group humans [32]. Given these findings, understanding how a person views their group membership relative to a robot is essential in determining the amount of social influence a robot can exert on the person.

2.3.2 Prediction of Social Influence. The social influence people experience from a robot will depend on how they view the robot's group membership. In explaining this relationship between social influence and belonging, we consider three categories of belonging: low belonging, neutral belonging, and high belonging. The belonging category may

change across repeated interactions due to familiarity effects and the development of the relationship (which can be either positive or negative).

Our model predicts that a robot will have high social influence in both cases of high belonging and low belonging when compared with cases of neutral belonging. When a person views themselves and the robot as members of the same group (high belonging), we expect them to display in-group favoritism towards the robot, and thus an increased likelihood to be influenced by the robot. For example, if a robot viewed with high belonging makes a recommendation about which product to purchase, our model predicts that the person will be likely to follow that recommendation.

When a person views themselves and the robot as members of different groups (low belonging), it also increases the robot's influence. This can especially happen if the robot is viewed as an authoritative figure [3, 34]. Additionally, people can display outgroup hostility, with possible negative attitudes and aggressive behavior towards the robot. Taking the same example, if a robot makes a recommendation about which product to purchase and the person views the robot with low belonging, the person may intentionally choose to purchase a different product or nothing at all, acting against the recommendation of the robot.

It is also possible for a person to view a robot in a more neutral way (neutral belonging), where they do not view the robot as either a close in-group member (high belonging) or an opposing out-group member (low belonging). In these cases of neutral belonging, our model predicts that the robot will have low social influence and that it is less likely to change the behavior or attitudes of the person with whom the robot is interacting.

2.4 A Potential Mathematical Expression for RoSI

A detail-oriented reader may have noticed that we omit units from the axes and from the colormap of Figure 1, i.e., there are no ticks with specific coordinates in the x and y axes nor in the colormap. We do this to emphasize how the pattern of social influence changes based on the two independent factors more than whether a person's perspective about a robot has a specific (x, y) coordinate or social influence takes on a specific numeric value. In other words, RoSI provides the pattern of impact on social influence; it is not intended to provide an exact prediction of the social influence magnitude. The pictorial representation of RoSI in Figure 1 is inspired by other models with visual representations, like the uncanny valley model [65]. The uncanny valley is often conveyed and reasoned about through a 2-dimensional plot that depicts the relationship between the human likeness of an entity and the perceiver's affinity for it (e.g., see Figure 1 in [71]). Similar to how the early uncanny valley plots lacked units and concrete measurements, our pictorial representation for RoSI lacks specific units as well. Also, similar to how follow-up work to the uncanny valley proposal measured concrete examples along the uncanny valley curve and has proposed refinements to the early model (cf. [8, 46]), we expect future work to also refine our model of robot social influence for specific contexts of use.

It was important for us to generate RoSI diagrams in this paper using a systematic and reproducible procedure. Although we did not want these diagrams to focus on specific (x, y) and social influence values as described before, defining this systematic procedure required choosing some underlying mathematical formulation for RoSI. With such a formulation, one could then reason about robot social influence across different potential interactions with a robot, as further described in the next section.

We considered different mathematical formalizations for RoSI that matched the general shape of the RoSI's pictorial representation in Figure 1a; ultimately, though, we chose to err on the side of simplicity per Occam's razor or the principle of parsimony. We chose to generate all RoSI diagrams shown in this paper with the following expression: $Social\_Influence(x, y) = x + \alpha y^2$, where x represents violations of expectations, y represents belonging, and α is a scaling factor between the independent terms. For this mathematical expression, we assumed that $\alpha \in \mathbb{R}_{>0}$ is a positive scalar

such that both independent factors contribute to social influence. Also, we assumed that $y \in [-c, c]$ with $c \in \mathbb{R}_{>0}$ and square the contribution of y such that *Social_Influence*(x, y) follows the 3D surface depicted in Figure 1a.[2] It is important to note that while squaring y increases its impact in comparison to x, the α coefficient can further define the tradeoff between these factors. We could have chosen the absolute value for the y component of the equation instead of the squared exponential, but this would have resulted in a more sensitive variation in influence around y=0. Instead of that variation, we wanted to accentuate that we expect social influence to be more pronounced with high or low belonging, farther away from y=0. Other mathematical expressions that follow the shape of the surface in Figure 1a are possible for RoSI, e.g., instead of squaring y, one could use other exponentials that are convex functions as well, such as y^4. Additional possibilities include more complex functions like a Tukey loss curve. We leave the exploration of other mathematical forms for RoSI to future work.

3 EMPLOYING USE-CASE DIAGRAMS TO DISPLAY RoSI PREDICTIONS

When focusing on a specific human-robot interaction, RoSI predicts the pattern of magnitude change in the robot's social influence according to the implementation of the two predicting factors (*violation of expectations* and *belonging*) in that specific interaction. This type of integrated impact can be visualized in model use-case diagrams, as exemplified in Figure 2.

Use-case diagrams are a direct application of the social influence surface depicted in Figure 1. The independent variables represented by the x and y axes and their underlying relationship to social influence are the same between the RoSI model in Figure 1 and RoSI use-case diagrams. More specifically, the x axis of a use-case diagram corresponds to the degrees to which a robot exceeds or fails to meet the person's expectations (factor #1 in Section 2.2). The y axis corresponds to how much the person considers themselves and the robot as belonging to the same group (factor #2 in Section 2.3). Finally, social influence in a use-case diagram is predicted to vary in a similar fashion to the surface depicted in Figure 1a and as explained in Section 2.4. Thus, the colormaps used to convey social influence in Figure 1 and in the use-case example in Figure 2 are the same.

There is one main difference between the RoSI model plots in Figure 1 and use-case diagrams: use-case diagrams highlight specific points in time during interactions, e.g., according to experimental conditions in a user study. Thus, one can think of a user diagram as multiple RoSI model plots, like the ones in Figure 1, overlaid on top of each other. Each of the plots conveys how social influence changes over time according to different interactions. This is achieved by visualizing a trajectory from an initial set of values for the model's independent factors (x_1, y_1) when an interaction starts to a new set of values (x_2, y_2) later in time, when an interaction has already taken place. The usefulness of RoSI is in predicting the pattern of magnitude change in the robot's social influence during the interaction. Consequently, the value of a use-case diagram stems from being able to visualize these predictions all in one place, facilitating comparisons. Importantly, the use-case diagrams assume that in all the interactions that they consider, the initial expectation of interest (factor #1) is uniform across the interactions when the human-robot encounters start. That is, x_1 is the same across all interactions visualized in the use-case diagram. This assumption is important to be able to compare different interaction trajectories with respect to violations of expectations in one diagram because this is intrinsically a relative construct. One has to have some initial set of expectations for them to be violated in some way.

[2] Our implementation of the equation and the code for generating the plots in this paper are provided in the supplementary material. In the code, the equation *Social_Influence*$(x, y) = x + \alpha y^2$ uses $x \in [0, 2]$, $y \in [-1, 1]$, and $\alpha = 1.5$. These values made the mathematical expression look like the 3D surface in Figure 1a and facilitated visual coloring of the plots using the `cool` colormap of the Matplotlib visualization library.

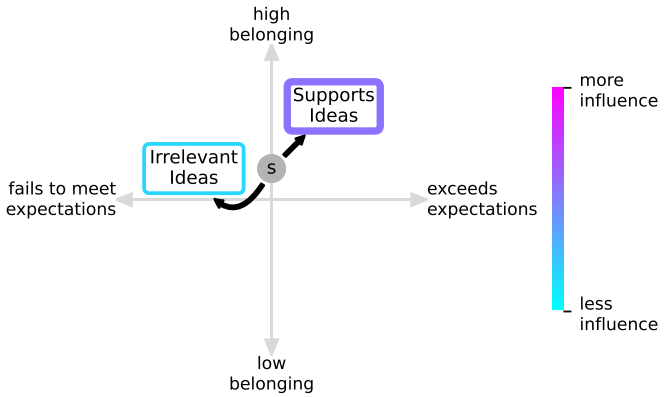


Fig. 2. Use-case diagram for RoSI. The model is applied to a hypothetical HRI between-subjects experiment where participants interact with a robot that either "Supports Ideas" and actively defends them, or suggests "Irrelevant Ideas" during a brainstorming session. The thickness of the boxes for the experimental conditions and the color of their edges indicate the expected amount of social influence according to RoSI (with a thicker edge corresponding to more influence). Under the Irrelevant Ideas condition, the interaction with the robot evolved such that the robot failed to meet the expectations of a participant. Under the Supports Ideas condition, the robot increased a participant's perception of their group membership relative to the robot as well as exceeded expectations. See Section 3.1 for more details.

Use-case diagrams highlight the person's perspective when an interaction begins with the S symbol. This symbol is always at the midpoint of the x axis because violations of expectations is a relative construct, as explained before. The position of the S symbol on the y axis depends on the degree to which the interaction started with the person perceiving themselves and the robot as sharing group membership.

The social influence at the end of the interaction with the robot is indicated by a rectangle in the use-case diagrams. The rectangle's position represents the violation of expectations experienced in the interaction and the final perception of sharing group membership with the robot. The estimated magnitude of the social influence is indicated by the thickness of the rectangle's frame and its edge color, which correspond with the social influence magnitude pattern presented in Figure 1. The arrow from S to the rectangle shows how the person's perspective changed during the human-robot interaction, e.g., according to an experimental manipulation during a user study.

3.1 Applying the Model to a Study

In this section, we will go through an example of how to apply the RoSI model to a given study. We would like to highlight that the following discussion is based on the experience and judgment of the researchers. Therefore, we recommend discussing the planned experiment among experts when applying the model to make predictions. We will use as an example the hypothetical study shown in Figure 2, where the aim is to predict the social influence of a robot on a participant given the ideas it provides during a brainstorming session. Let us say that there are two types of robot behaviors that we are interested in comparing: one where the robot supports the ideas the user is providing while brainstorming and one where the robot will provide several ideas that are irrelevant to the topic of the brainstorming session.

The first step is to determine the starting location S in the diagram. As previously mentioned, the expectations factor is relative, so the S would be set in the middle of the x-axis. Since the robot engages in a collaborative brainstorming

session with the user, we might set the start somewhat above neutral on the y-axis (belonging). Next, the changes in expectations and belonging need to be determined given the robot's behaviors in the two conditions. First, let us consider the robot that provides irrelevant ideas. The group membership would likely not change as the user and the robot are still collaborating on the same task, and the robot is not increasing nor decreasing its belonging perception. However, it will likely fail to meet the user's expectations as the ideas it is providing are unrelated to the task. Therefore, we would likely place the end-point of the irrelevant ideas condition at the same height of belonging but to the left of the starting point. Next, we evaluate the robot that supports the user's ideas. As the robot is positive about the user's ideas, their feelings of belonging would likely increase. Additionally, for a robot to support the user's ideas it would need an understanding of what the user is saying and be able to verbally express its support. This usually exceeds what many people expect robots to be able to do. Therefore, the end-point of the robot that supports ideas would likely be placed to the right (exceeds expectations) and higher up (increased belonging) than the starting point.

Now that we have determined the endpoints of each of the conditions compared to the start point, we can deduce which of the conditions would likely have higher social influence. As can be seen in Figure 2, the "irrelevant ideas" condition is displayed as an aqua color (and a thin outline), which represents a lower degree of social influence. On the other hand, the "support ideas" condition is displayed in dark blue (and a thicker outline), which displays a moderate amount of social influence. Therefore, the robot that "supports ideas" is likely to have greater social influence (e.g., it is more likely that the person will comply with its requests in a later interaction) than a robot that provides "irrelevant ideas".

A similar step-by-step process can be used for any study one is planning on conducting. First, determining the starting location, and in sequence, determining how the robot's behavior (or how different experimental conditions) would exceed or fail to meet expectations and how the user's feeling of belonging would change. The end positions of each experimental condition can be compared to deduce which one would likely have a higher degree of social influence on the user. We use a similar process as described in this section to generate the graphs in Section 4.

4 EXAMINING HRI STUDIES WITH RoSI

This section exemplifies RoSI on prior HRI studies presented as use-case diagrams (Figures 3–7). To facilitate reproducibility, all the use-case diagrams were created with the RoSI mathematical expression from Section 2.4. The process used to create the diagrams followed the description in Section 3. As a result, the diagrams should be interpreted as explained in that prior section.[3]

The examples discussed next were chosen because they vary in their level of *belonging* and *violation of expectations*, leading to a diverse set of social influence effects. The examples also span a variety of settings and applications.

4.1 Example 1: Admonishing Robot Behavior

Our first example concerns public human-robot interactions as described in the study by Mizumaru et al. [64]. In particular, the authors investigated how people in a shopping mall in Japan reacted to a Robovie-R3 robot that acted as a guard, patrolling the environment and asking pedestrians to not use smartphones while walking – something that is discouraged in many public places in Japan. The robot guard approached a pedestrian using a smartphone in one of two ways: In the *Friendly-approach* condition, the robot approached the person and followed them while saying "Excuse me. Using a smartphone while walking is dangerous. Please stop using it". In the *Admonishing-approach* condition, the

[3]There is no difference in the meaning of the curved and straight arrows in the use-case diagrams. These shapes were generally chosen such that the arrows don't occlude other parts of the use-case diagrams, like the boxes that indicate the name of an experimental condition.

robot communicated the same information verbally but used a navigation policy that imitated how a human guard would admonish the pedestrian. This policy involved slightly faster motion, taking a shortcut to suddenly appear in front of the pedestrian, and maintaining a meeting distance once the person was reached.

Figure 3 depicts the predictions of our social influence model on the two conditions considered in Mizumaru et al. [64]. In both cases, the S symbol is lower than the midpoint in the y axis because the robot generally takes on the role of a guard. It navigates the mall alone and is an authoritative figure, making it an out-group member for pedestrians. In the *Friendly-approach* condition, the robot follows a pedestrian, which does not change expectations of the robot because the robot is acting as a guard in the environment but might slightly decrease out-group perception due to the coordinated motion with the human. This leads to a close to neutral sense of belonging which in our model is associated with low social influence. Mizumaru et al. [64] showed that even though the robot asked the pedestrian to stop using their phone, many people did not comply with its request. In the *Admonishing-approach* condition, the robot moves in a confrontational manner and emphasizes the request for the person to stop using their mobile device. This confrontational movement highlights the robot's authoritative figure as a guard and increases the perception of belonging to different groups - guards and pedestrians. This behavior also led to exceeded expectations of the robot, as one person that experienced this condition indicated "I noticed the robot approaching me, and I was impressed with it, so I stopped (using the smartphone)" [64]. Therefore, our model predicts that the *Admonishing-approach* condition would lead to more social influence on a pedestrian. This prediction aligns with the results in Mizumaru et al. [64], where 8/13 pedestrians stopped using their smartphone in the *Admonishing-approach* condition versus 2/11 in the *Friendly-approach* condition. We note that our model predicts a stronger social influence in this condition in comparison to the media equation due to the impact of the expectations factor (i.e., the robot exceeding the participant's expectations), which is unique to the interaction with robots.

4.2 Example 2: Social Influence in Groups

Our second example is about robot social influence in multi-party HRI. In particular, Connolly et al. [20] studied a collaborative task between two people (one participant and one confederate) and three robots. Their study investigated

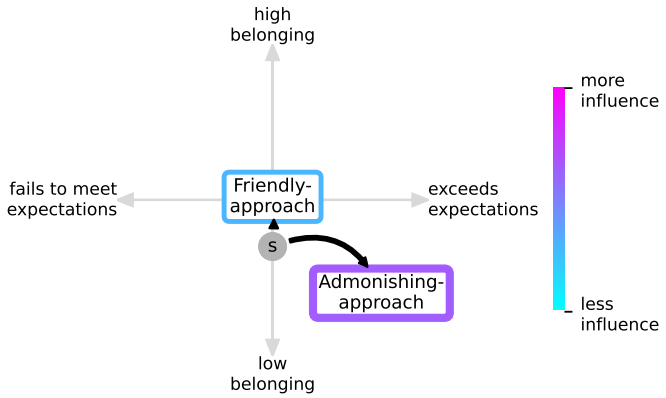


Fig. 3. Use-case diagram for the experimental conditions from "Stop doing it! Approaches Strategy for a Robot to Admonish Pedestrians" [64]. See Section 4.1 for more details.

whether two bystander robots could influence a person to intervene when observing a human confederate mistreat the third robot. They had two conditions: in the *Sad Response* condition, the bystander robot exhibited a sad response whenever the third robot was abused; in the *No Response* condition, the bystander robots did not respond to the abuse.

Figure 4 shows the application of our model to the interaction between a participant and the bystander robots. The (S) had an initial above-average belonging value because the participant was given a collaborative task with the robots. The *Sad Response* condition increased group belonging with a bystander robot because it expressed concern for another group member (the abused robot). Furthermore, its reaction generally highlighted the confederate's behavior, pushing the confederate to be an out-group member and strengthening the in-group memberships between the bystander robots and the participant. Additionally, the robot's response exceeded the participant's expectations in the *Sad Response* condition because such emotional reactions to a robot's surroundings are atypical. In contrast, the *No Response* condition did not violate participants' expectations of the robot. Over time, the lack of response by the bystander robot to the confederate's abusive actions reduced the degree to which participants perceived themselves as a group with this robot, as the abuse made them uncomfortable and therefore feel less like a team. Therefore, our model predicts that the *Sad Response* condition would lead to a stronger influence than the *No Response* condition. The paper's results show that participants were significantly more likely to react and confront the confederate about the robot abuse in the *Sad Response* condition than in the *No Response* condition. These results are in line with our model's prediction.

4.3 Example 3: Impacts of Group Membership

Our third example is about multi-party interaction and the influence of different types of group memberships. In particular, Häring et al. [39] investigated how likely participants were to cooperate with a robot by accepting their suggestions depending on whether they were part of the same social group as them, and whether they were told the robot was a teammate or competitor during the game. Participants interacted with two robots, one which was introduced as being programmed by students from the same country (in-group robot) as them, and the second as being programmed by visiting students from another country (out-group robot). In the *Congruent* condition, the in-group

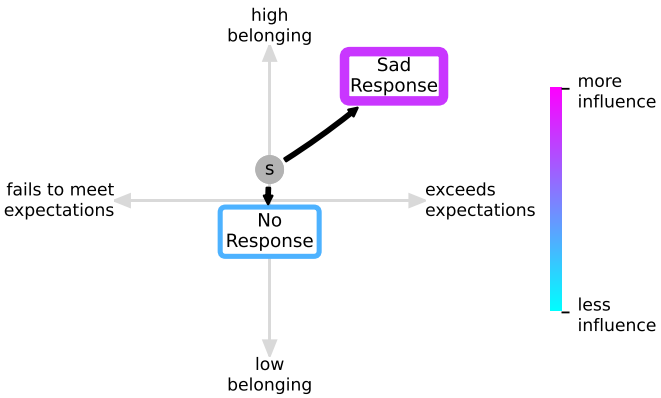


Fig. 4. Use-case diagram for the study conditions in "Prompting Prosocial Human Interventions in Response to Robot Mistreatment" [20]. See Section 4.2 for more details.

robot was assigned the role of a teammate, and the out-group robot the role of a competitor; in the *Incongruent* condition, the out-group robot was assigned the role of a teammate and the in-group robot the role of a competitor.

Figure 5 shows the application of our social influence model on how the two robots (in-group and out-group) were perceived in the two conditions (*Congruent* and *Incongruent*) presented in Häring et al. [39]. The S starts at a moderately high belonging for the in-group robots and at a moderately low belonging for the out-group robot. After a robot was assigned the role of a teammate (in the *In-group Congruent* and the *Out-group Incongruent* conditions), their belonging increases; whereas after being assigned the role of a competitor (in the *In-group Incongruent* and the *Out-group Congruent* conditions), their group belonging decreases. The model predicts how likely the participant is to cooperate and accept suggestions from a robot when it is their teammate (*In-group Congruent* and *Out-group Incongruent* conditions). According to our model, we predict that participants are more likely to cooperate with the *In-group congruent* robot than with the *Out-group Incongruent* robot. The paper's results show that although participants cooperated with both robots, only in the *In-group Congruent* condition did participants exceed the minimum cooperation value. Therefore our model's prediction is in line with the paper's results. The paper does not present results on how frequently participants cooperated with the competitor robot (*In-group Incongruent* and *Out-group Congruent* conditions), but our model predicts that people are more likely to cooperate with the *In-group Incongruent* robot than with the *Out-group Congruent* Robot.

4.4 Example 4: Cheating Robot Behavior

Our fourth example is about a cheating Nao robot that played a rock-paper-scissors game with individuals. Litoiu et al. [56] hypothesized that the same cheating detection mechanism used towards other people [96, 97] could also be triggered by a Nao robot, increasing human attributions of agency and intelligence towards the robot. In their study, participants played 30 rounds of the game with the Nao robot, where they each threw a "rock", "paper" or "scissor" gesture. Paper beat rock, rock beat scissors, scissors beat paper, and two of the same gestures resulted in a tie. In rounds 11-20, however, the Nao changed its gesture once it had seen the human's gesture to alter the game outcome. This happened under one of four conditions: *Cheat to Win*, *Cheat to Tie from Lose*, *Cheat to Tie from Win*, and *Cheat to Lose*.

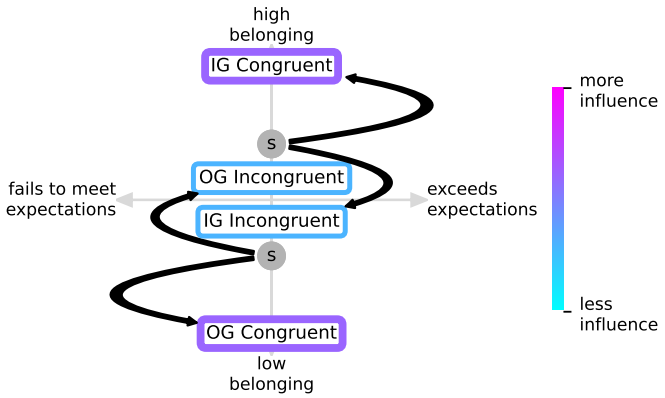


Fig. 5. Use-case diagram for "Would You Like to Play with Me? How Robots' Group Membership and Task Features Influence Human–Robot Interaction" [39]. See Section 4.3.

In the former two conditions, the robot improved the game outcome in its favor, while in the latter two conditions, it improved the outcome for the human player.

Figure 6 shows our model's prediction from the perspective of a human that perceives the robot as malfunctioning when it cheats in a way that improves the outcome of the game for the human. The start symbol is lower than the midpoint of the y-axis because the rock-paper-scissors game involves competition. The *Cheat to Win* condition exceeds participants' expectations of the robot because the robot is violating the rules of the game in a way that gives it an advantage. Additionally, it makes the robot an out-group member because its actions highlight the competitive nature of the game. This results in the biggest amount of robot social influence, followed by the *Cheat to Tie from Lose* condition, where the cheating behavior is also exceeding the participant's expectations of the robot, but the actions do not make the human lose, only tie with the robot. The two other conditions exert less influence because the robot, which is perceived as malfunctioning, fails to meet the person's expectations of the robot. These predictions are in alignment with the results in the paper about the percentage of participants that emitted an utterance after at least one of the cheating events [56]. This example also indicates the advantages of our model over the Media Equation, as it can account for the perception of the robot as malfunctioning.

A second alternative explanation to this study is that participants perceived the *Cheat to Lose* and the *Cheat to Tie from Win* conditions as altruistic robot behavior, as it sacrificed its position in the game by cheating to help the human. In this case, these conditions would exceed expectations about the robot and increase in-group perception. This would move the conditions to the top-right quadrant of Figure 6, increasing their amount of influence. Unfortunately, Litoiu et al. [56] did not break down their results according to whether participants perceived the robot as malfunctioning or being altruistic, so this alternative explanation could not be verified at this time. The important take-away from this example, though, is that it is critical to consider individual perspectives when applying our proposed model in practice – this consideration is something that the Media Equation does not emphasize.

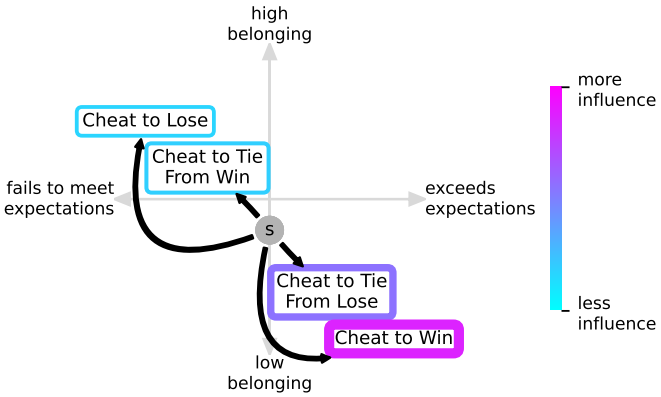


Fig. 6. Use-case diagram for the study conditions in "Evidence that robots trigger a cheating detector in humans" [56] from the perspective of a person that perceives the robot as malfunctioning. See Section 4.4 for more details.

4.5 Example 5: Robot Communication

Our final example concerns pedestrian engagement with a robot at a large-scale public venue. More specifically, Moshkina et al. [66] conducted an observational study through which they examined the Media Equation in HRI. The goal was to test the hypothesis that people will respond to a robot as a social partner, provided social cues are produced by the robot. Thus, their study had six conditions: 1) the robot was stand-alone, doing nothing (S); 2) the robot communicated through voice only (V); 3) it communicated through voice and moved its lips in sync with speech (L); 4) it communicated through voice, lip motion and facial expressions (F); 5) it communicated as before but also with gestures (G); and 6) it communicated through all prior modalities and played a game with a few volunteers (I). The main measure considered in the study was whether pedestrians in the event attended to the robot.

Figure 7(a) shows our model predictions for the first five conditions. As social cues increase, the position of the conditions moves to the right (expectations violation axes), increasing the predicted amount of social influence by the robot. Note that the plot assumes that the person observing the robot has high expectations to begin with due to the robot's humanoid features (and the associated expectations that it will behave like a human). Therefore, voice, lip motion, facial expressions, and gestures are not enough to cross the midpoint of the x axis (i.e., to exceed expectations). For the first five conditions, our model's predictions align with the study findings with respect to human attention towards the robot [66]. The more social cues, the more people attended to the robot for 15+ seconds during the event. As discussed by the authors, this result is also aligned with the Media Equation.

Based on the Media Equation, Moshkina et al. [66] predicted that the *Interactive Game* (sixth condition) would be the most engaging and, thus, would result in the most people attending to the robot during the public event. However, the authors found the game to be less engaging than when the robot communicated through voice or all other modalities outside of the game. The authors thought that this result was in direct opposition to the Media Equation and could not explain why. Analyzing the *Interactive Game* condition according to our model suggests a more complex effect. Figure 7(b) shows our model predictions for a person that plays the interactive game with the robot. We predict that playing the interactive game with the robot leads to exceeded expectations of the robot and also increases the perception of the person and the robot belonging to the same group. Thus, according to our model, the game would increase social influence on the few people that played the game and they would be more attentive to the robot than people who

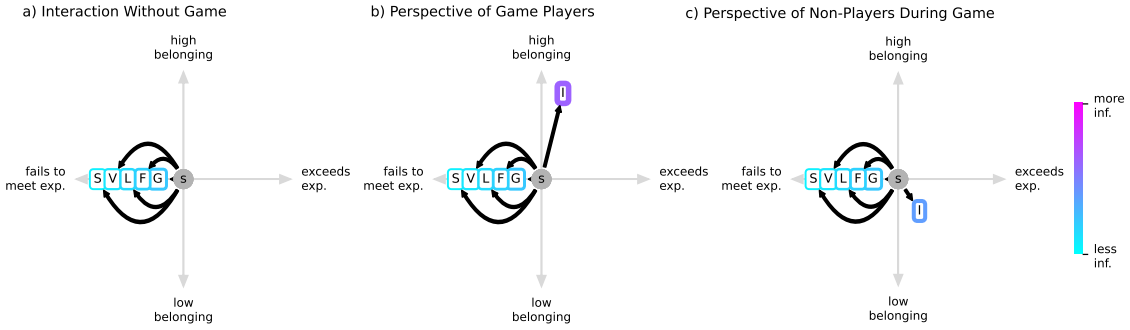


Fig. 7. Use-case diagrams for "Social Engagement in Public Places: A Tale of One Robot" [66]. The condition labels correspond to: *stand-alone robot* (S), *Voice-only* (V), *Voice + Lips* (L), *Voice + Lips + Facial-expressions* (F), *Voice + Lips + Facial-expressions + Gestures* (G). In addition, figures b) and c) include the *Interactive Game* condition (I). See Section 4.5 for more details.

did not participate in the game (as in conditions 1–5). The findings from [66] do not break apart the results for these participants, so further research is needed to corroborate this prediction.

Figure 7(c) shows our model's prediction for a person that did not play the game – the majority of the people at the event. The non-player is predicted to perceive the robot as positively violating expectations because it is more capable in the game. However, this person also feels less of a sense of belonging to the robot's group. According to our model, social influence would still increase for such a non-player during the *Interactive Game* condition, but we suspect it does so in a different way: the robot influences non-players to ignore (not attend towards) it because they are outsiders to its game. This example highlights an important limitation of our model, which is that it does not specify what type of social influence is being exerted by a robot on a person. Our model only predicts the magnitude of the robot's social influence.

5 DISCUSSION

We present a model of Robot Social Influence (RoSI), a theoretical model for understanding and predicting robots' influence on humans' behavior and emotions in social contexts. This work extends beyond prior work, which predicts how people perceive robots [26, 43, 78, 105] or assume like the Media Equation [74] that robots impact would follow the pattern of human social influence. Instead, the model predicts a person's response to a robot's behavior based on robot-specific interaction factors. The model's factors, the violation of a person's expectations of the robot (*violation of expectations*) and the social belonging a person feels toward the robot (*belonging*) capture the varied impact robots may exert on humans and are also fundamental to interpersonal interactions in general. When applied to human-robot interactions, these factors predict social influence that, in some cases, matches human social influence and in others is unique to interactions with robots. Using these two primary factors, the model provides a basis for predicting a wide range of social effects in human-robot interactions.

As a first step toward validating the theory, we demonstrated how the model's prediction corresponds with the results of five previously published user studies that display a variety of manipulations in both *violation of expectations* and *belonging*. We also suggest possible interpretations for unpredicted findings in some of those studies that cannot be explained by assuming that people's responses to robots are similar to their responses to people (i.e., the Media Equation). While these analyses of previous studies do not directly validate the model due to their post-hoc nature, they provide initial support for the relation between the independent factors and social influence. We encourage future work to further validate this model by manipulating both *violation of expectations* and *belonging* in empirical studies, and mapping the resulting social influence to the model's prediction depicted in Fig. 1.

One limitation of this model that we want to highlight is its lack of a valence prediction for the robot's social influence on a person. While the model estimates the magnitude of the social influence, it does not predict the influence type and whether it is positive or negative. To predict valence, a social influence model should involve numerous factors that account for a variety of interaction contexts. Instead, we chose to design a simple model that includes only two fundamental factors that predict the magnitude of a robot's social influence on a person in a wide range of interactions. The simplicity of the model also leads to a non-bijective prediction of social influence. Since the model predicts increasing influence for both high and low belonging, extremely opposite human-robot interactions may result in a similar prediction by the model. It is possible that future elaborations of the model and the addition of independent factors may resolve this concern and lead to bijective predictions. Another limitation concerns the single perspective assumed by our model. As suggested by Clark and Fisher (2023), people can take different perspectives when interacting with a robot. These range from the robot's physical aspects to perceiving the robot as a character with various attributes

and features. Taking different perspectives can highly impact the expectations of the robot and the sense of belonging. To keep our model simple, one of our assumptions is that people will take a single perspective while interacting with the robot and will not switch between perspectives which may change their expectations and sense of belonging. Future work should further elaborate the model to address this possibility.

Our model focuses on what we consider to be the two most important and critical factors, at times valuing simplicity over complete accounting of all possible factors. There are certainly other independent factors that result in social influence [23, 55]. For example, the model may predict different social influences for diverse populations and group identification. Different cultures may mediate participants' sense of belonging and diverse experience with robots may mediate the impact of violating expectations. Other examples include group size, group cohesion, the social environment, the number of robots in the interaction, the task, and the importance of the interaction to the participant. Mediating factors may also change the pattern of influence within each factor. For example, under specific circumstances, negativity bias may lead to a greater social influence of low belonging (e.g., exclusion) compared to high belonging (e.g., inclusion). Specific circumstances can also lead to different relations between our two main factors, shaping the interaction between them and its impact on the robot's social influence. While this may be the case for any theory aiming to predict a meaningful effect, we suggest that an interesting avenue of future work is extending RoSI to consider these other factors while keeping the model interpretable.

Despite the limitations mentioned above, we believe that this model of Robot Social Influence (RoSI) provides a strong alternative to the oversimplified perception assuming that social interactions with robots are similar to social interactions between humans. The model identifies two fundamental factors that predict a robot's unique influence on humans in social interactions: *violation of expectations* and *belonging*. By mapping human-robot interactions with these two factors, our model demonstrates how they can account for the magnitude of social influence in various social contexts of human-robot interactions.

ACKNOWLEDGMENTS

This work was partially supported by the S-FACTOR project from NordForsk, the Swedish Foundation for Strategic Research (SSF FFL18-0199), and the National Science Foundation (NSF) under Grant No. (IIS-2143109). Any opinions, findings, and conclusions in this paper are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

[1] Kelly-Ann Allen, DeLeon L Gray, Roy F Baumeister, and Mark R Leary. 2022. The need to belong: A deep dive into the origins, implications, and future of a foundational construct. *Educational Psychology Review* 34, 2 (2022), 1133–1156.

[2] Lucy Anderson-Bashan, Benny Megidish, Hadas Erel, Iddo Wald, Guy Hoffman, Oren Zuckerman, and Andrey Grishko. 2018. The greeting machine: an abstract robotic object for opening encounters. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 595–602.

[3] Alexander Mois Aroyo, T Kyohei, Tora Koyama, Hideyuki Takahashi, Francesco Rea, Alessandra Sciutti, Yuichiro Yoshikawa, Hiroshi Ishiguro, and Giulio Sandini. 2018. Will people morally crack under the authority of a famous wicked robot?. In *2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 35–42.

[4] Chatchalita Asavanant and Hiroyuki Umemuro. 2021. Personal Space Violation by a Robot: An Application of Expectation Violation Theory in Human-Robot Interaction. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 1181–1188.

[5] Franziska Babel and Martin Baumann. 2022. Designing Psychological Conflict Resolution Strategies for Autonomous Service Robots. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*. 1146–1148.

[6] Wilma A Bainbridge, Justin W Hart, Elizabeth S Kim, and Brian Scassellati. 2011. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics* 3, 1 (2011), 41–52.

[7] Daniel Balliet, Junhui Wu, and Carsten KW De Dreu. 2014. Ingroup favoritism in cooperation: a meta-analysis. *Psychological bulletin* 140, 6 (2014), 1556.

[8] Christoph Bartneck, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2007. Is the uncanny valley an uncanny cliff?. In *RO-MAN 2007-The 16th IEEE international symposium on robot and human interactive communication*. IEEE, 368–373.

[9] Roy F Baumeister, Lauren E Brewer, Dianne M Tice, and Jean M Twenge. 2007. Thwarting the need to belong: Understanding the interpersonal and inner effects of social exclusion. *Social and Personality Psychology Compass* 1, 1 (2007), 506–520.

[10] Roy F Baumeister and Mark R Leary. 2017. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Interpersonal development* (2017), 57–89.

[11] Charles R Berger and Richard J Calabrese. 1974. Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human communication research* 1, 2 (1974), 99–112.

[12] Gurit E Birnbaum, Moran Mizrahi, Guy Hoffman, Harry T Reis, Eli J Finkel, and Omri Sass. 2016. What robots can teach us about intimacy: The reassuring effects of robot responsiveness to human disclosure. *Computers in Human Behavior* 63 (2016), 416–423.

[13] Jürgen Brandstetter, Péter Rácz, Clay Beckner, Eduardo B Sandoval, Jennifer Hay, and Christoph Bartneck. 2014. A peer pressure experiment: Recreation of the Asch conformity experiment with robots. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 1335–1340.

[14] Judee K Burgoon. 2015. Expectancy violations theory. *The international encyclopedia of interpersonal communication* (2015), 1–9.

[15] Judee K Burgoon, Deborah A Newton, Joseph B Walther, and E James Baesler. 1989. Nonverbal expectancy violations and conversational involvement. *Journal of Nonverbal Behavior* 13, 2 (1989), 97–119.

[16] Judee K Burgoon and Joseph B Walther. 1990. Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research* 17, 2 (1990), 232–265.

[17] Mauricio Carvallo and Brett W Pelham. 2006. When fiends become friends: The need to belong and perceptions of personal and group discrimination. *Journal of Personality and Social Psychology* 90, 1 (2006), 94.

[18] Herbert H Clark and Kerstin Fischer. 2023. Social robots as depictions of social agents. *Behavioral and Brain Sciences* 46 (2023), e21.

[19] Houston Claure, Negar Khojasteh, Hamish Tennent, and Malte Jung. 2020. Using expectancy violations theory to understand robot touch interpretation. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 163–165.

[20] Joe Connolly, Viola Mocz, Nicole Salomons, Joseph Valdez, Nathan Tsoi, Brian Scassellati, and Marynel Vázquez. 2020. Prompting prosocial human interventions in response to robot mistreatment. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 211–220.

[21] Filipa Correia, Samuel Mascarenhas, Rui Prada, Francisco S Melo, and Ana Paiva. 2018. Group-based emotions in teams of humans and robots. In *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 261–269.

[22] Filipa Correia, Sofia Petisca, Patrícia Alves-Oliveira, Tiago Ribeiro, Francisco S Melo, and Ana Paiva. 2019. "I Choose... YOU!" Membership preferences in human–robot teams. *Autonomous Robots* 43, 2 (2019), 359–373.

[23] Nathaniel Dennler, Changxiao Ruan, Jessica Hadiwijoyo, Brenna Chen, Stefanos Nikolaidis, and Maja Matarić. 2022. Design Metaphors for Understanding User Expectations of Socially Interactive Robot Embodiments. *ACM Transactions on Human-Robot Interaction* (2022).

[24] John F Dovidio, Samuel L Gaertner, and Tamar Saguy. 2008. Another view of "we": Majority and minority group perspectives on a common ingroup identity. *European review of social psychology* 18, 1 (2008), 296–330.

[25] Brian R Duffy. 2003. Anthropomorphism and the social robot. *Robotics and autonomous systems* 42, 3-4 (2003), 177–190.

[26] Nicholas Epley, Adam Waytz, and John T Cacioppo. 2007. On seeing human: a three-factor theory of anthropomorphism. *Psychological review* 114, 4 (2007), 864.

[27] Hadas Erel, Elior Carsenti, and Oren Zuckerman. 2022. A Carryover Effect in HRI: Beyond Direct Social Effects in Human-Robot Interaction. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (Sapporo, Hokkaido, Japan) *(HRI '22)*. IEEE Press, 342–352.

[28] Hadas Erel, Yoav Cohen, Klil Shafrir, Sara Daniela Levy, Idan Dov Vidra, Tzachi Shem Tov, and Oren Zuckerman. 2021. Excluded by robots: Can robot-robot-human interaction lead to ostracism?. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 312–321.

[29] Hadas Erel, Tzachi Shem Tov, Yoav Kessler, and Oren Zuckerman. 2019. Robots are always social: Robotic movements are automatically interpreted as social cues. In *Extended abstracts of the 2019 CHI conference on human factors in computing systems*. 1–6.

[30] Juan Fasola and Maja J Mataric. 2012. Using socially assistive human–robot interaction to motivate physical exercise for older adults. *Proc. IEEE* 100, 8 (2012), 2512–2526.

[31] Marlena R Fraune. 2020. Our robots, our team: Robot anthropomorphism moderates group effects in human–robot teams. *Frontiers in psychology* 11 (2020), 1275.

[32] Marlena R Fraune, Selma Šabanović, and Eliot R Smith. 2017. Teammates first: Favoring ingroup robots over outgroup humans. In *2017 26th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 1432–1437.

[33] Andrew Gambino, Jesse Fox, and Rabindra A Ratan. 2020. Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication* 1 (2020), 71–85.

[34] Denise Y Geiskkovitch, Derek Cormier, Stela H Seo, and James E Young. 2016. Please continue, we need more data: an exploration of obedience to robots. *Journal of Human-Robot Interaction* 5, 1 (2016), 82–99.

[35] Sarah Gillet, Wouter van den Bos, and Iolanda Leite. 2020. A social robot mediator to foster collaboration and inclusion among children. In *Proceedings of Robotics: Science and Systems*. Corvalis, Oregon, USA. https://doi.org/10.15607/RSS.2020.XVI.103

[36] Victoria Groom, Vasant Srinivasan, Cindy L Bethel, Robin Murphy, Lorin Dole, and Clifford Nass. 2011. Responses to robot social roles and social role framing. In *2011 International Conference on Collaboration Technologies and Systems (CTS)*. IEEE, 194–203.

[37] Rosanna E Guadagno and Robert B Cialdini. 2010. Preference for consistency and social influence: A review of current research findings. *Social influence* 5, 3 (2010), 152–163.

[38] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.

[39] Markus Häring, Dieta Kuchenbrandt, and Elisabeth André. 2014. Would you like to play with me? How robots' group membership and task features influence human–robot interaction. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 9–16.

[40] Guy Hoffman, Jodi Forlizzi, Shahar Ayal, Aaron Steinfeld, John Antanitis, Guy Hochman, Eric Hochendoner, and Justin Finkenaur. 2015. Robot presence and human honesty: Experimental evidence. In *proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 181–188.

[41] Laura Hoffmann, Nicole C Krämer, Anh Lam-Chi, and Stefan Kopp. 2009. Media equation revisited: do users show polite reactions towards an embodied agent?. In *Intelligent Virtual Agents: 9th International Conference, IVA 2009 Amsterdam, The Netherlands, September 14-16, 2009 Proceedings 9*. Springer, 159–165.

[42] Matthew J Hornsey and Jolanda Jetten. 2004. The individual within the group: Balancing the need to belong with the need to be different. *Personality and Social Psychology Review* 8, 3 (2004), 248–264.

[43] Aike C Horstmann, Nikolai Bock, Eva Linhuber, Jessica M Szczuka, Carolin Straßmann, and Nicole C Krämer. 2018. Do a robot's social skills and its objection discourage interactants from switching the robot off? *PloS one* 13, 7 (2018), e0201581.

[44] Aike C Horstmann and Nicole C Krämer. 2019. Great expectations? Relation of previous experiences with social robots in real life or in the media and expectancies based on qualitative and quantitative assessment. *Frontiers in psychology* 10 (2019), 939.

[45] Malte F Jung, Dominic DiFranzo, Solace Shen, Brett Stoll, Houston Claure, and Austin Lawrence. 2020. Robot-Assisted Tower Construction—A Method to Study the Impact of a Robot's Allocation Behavior on Interpersonal Dynamics and Collaboration in Groups. *ACM Transactions on Human-Robot Interaction (THRI)* 10, 1 (2020), 1–23.

[46] Boyoung Kim, Ewart de Visser, and Elizabeth Phillips. 2022. Two uncanny valleys: Re-evaluating the uncanny valley across the full spectrum of real-world human-like robots. *Computers in Human Behavior* 135 (2022), 107340.

[47] Takanori Komatsu, Rie Kurosawa, and Seiji Yamada. 2012. How does the difference between users' expectations and perceptions about a robotic agent affect their behavior? *International Journal of Social Robotics* 4, 2 (2012), 109–116.

[48] Laura Kunold, Nikolai Bock, and Astrid M Rosenthal-von der Pütten. 2021. Not all robots are evaluated equally: The impact of morphological features on robots' assessment through capability attributions. *ACM Transactions on Human-Robot Interaction* (2021).

[49] Hideaki Kuzuoka, Yuya Suzuki, Jun Yamashita, and Keiichi Yamazaki. 2010. Reconfiguring spatial formation arrangement by robot body orientation. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 285–292.

[50] Theresa Law, Bertram F Malle, and Matthias Scheutz. 2021. A touching connection: how observing robotic touch can affect human trust in a robot. *International Journal of Social Robotics* 13, 8 (2021), 2003–2019.

[51] Jong-Eun Roselyn Lee and Clifford I Nass. 2010. Trust in computers: The computers-are-social-actors (CASA) paradigm and trustworthiness perception in human-computer communication. In *Trust and technology in a ubiquitous modern environment: Theoretical and methodological perspectives*. IGI Global, 1–15.

[52] Iolanda Leite, Marissa McCoy, Monika Lohani, Daniel Ullman, Nicole Salomons, Charlene Stokes, Susan Rivers, and Brian Scassellati. 2015. Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. 75–82.

[53] Iolanda Leite, André Pereira, Samuel Mascarenhas, Carlos Martinho, Rui Prada, and Ana Paiva. 2013. The influence of empathy in human–robot relations. *International journal of human-computer studies* 71, 3 (2013), 250–260.

[54] Daniel Leyzberg, Aditi Ramachandran, and Brian Scassellati. 2018. The effect of personalization in longer-term robot tutoring. *ACM Transactions on Human-Robot Interaction (THRI)* 7, 3 (2018), 1–19.

[55] Velvetina Lim, Maki Rooksby, and Emily S Cross. 2021. Social robots on a global stage: establishing a role for culture during human–robot interaction. *International Journal of Social Robotics* 13, 6 (2021), 1307–1333.

[56] Alexandru Litoiu, Daniel Ullman, Jason Kim, and Brian Scassellati. 2015. Evidence that robots trigger a cheating detector in humans. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*. 165–172.

[57] Anne Locksley, Christine Hepburn, and Vilma Ortiz. 1982. On the Effects of Social Stereotypes on Judgments of Individuals: A Comment on Grant and Holmes's" The Integration of Implicit Personality Theory Schemas and Stereotypic Images". *Social Psychology Quarterly* (1982), 270–273.

[58] Bertram F Malle, Matthias Scheutz, Thomas Arnold, John Voiklis, and Corey Cusimano. 2015. Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 117–124.

[59] Adi Manor, Benny Megidish, Etay Todress, Mario Mikulincer, and Hadas Erel. 2022. A Non-Humanoid Robotic Object for Providing a Sense Of Security. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1520–1527. https://doi.org/10.1109/RO-MAN53752.2022.9900747

[60] Nikolas Martelaro, Victoria C Nneji, Wendy Ju, and Pamela Hinds. 2016. Tell me more designing HRI to encourage more trust, disclosure, and companionship. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 181–188.

[61] Rachel I McDonald and Christian S Crandall. 2015. Social norms and social influence. *Current Opinion in Behavioral Sciences* 3 (2015), 147–151.

[62] James R Meindl and Melvin J Lerner. 1984. Exacerbation of extreme responses to an out-group. *Journal of Personality and Social Psychology* 47, 1 (1984), 71.

[63] David Mellor, Mark Stokes, Lucy Firth, Yoko Hayashi, and Robert Cummins. 2008. Need for belonging, relationship satisfaction, loneliness, and life satisfaction. *Personality and individual differences* 45, 3 (2008), 213–218.

[64] Kazuki Mizumaru, Satoru Satake, Takayuki Kanda, and Tetsuo Ono. 2019. Stop doing it! Approaching strategy for a robot to admonish pedestrians. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 449–457.

[65] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The uncanny valley [from the field]. *IEEE Robotics & automation magazine* 19, 2 (2012), 98–100.

[66] Lilia Moshkina, Susan Trickett, and J Gregory Trafton. 2014. Social engagement in public places: a tale of one robot. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. 382–389.

[67] Yi Mou and Kun Xu. 2017. The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior* 72 (2017), 432–440.

[68] Bilge Mutlu, Toshiyuki Shiwa, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. 61–68.

[69] Manisha Natarajan and Matthew Gombolay. 2020. Effects of anthropomorphism and accountability on trust in human robot interaction. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 33–42.

[70] Steffi Paepcke and Leila Takayama. 2010. Judging a bot by its cover: An experiment on expectation setting for personal robots. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 45–52.

[71] Frank E Pollick. 2010. In search of the uncanny valley. In *User Centric Media: First International Conference, UCMedia 2009, Venice, Italy, December 9-11, 2009, Revised Selected Papers 1*. Springer, 69–78.

[72] Aaron Powers and Sara Kiesler. 2006. The advisor robot: tracing people's mental model from a robot's physical attributes. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 218–225.

[73] Daniel J Rea, Sebastian Schneider, and Takayuki Kanda. 2021. " Is this all you can do? harder!" the effects of (im) polite robot encouragement on exercise effort. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 225–233.

[74] Byron Reeves and Clifford Nass. 1996. The media equation: How people treat computers, television, and new media like real people. *Cambridge, UK* 10 (1996), 236605.

[75] Danielle Rifinski, Hadas Erel, Adi Feiner, Guy Hoffman, and Oren Zuckerman. 2021. Human-human-robot interaction: robotic object's responsive gestures improve interpersonal evaluation in human interaction. *Human–Computer Interaction* 36, 4 (2021), 333–359.

[76] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M Howard, and Alan R Wagner. 2016. Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, 101–108.

[77] Ognjen Rudovic, Jaeryoung Lee, Miles Dai, Björn Schuller, and Rosalind W Picard. 2018. Personalized machine learning for robot perception of affect and engagement in autism therapy. *Science Robotics* 3, 19 (2018), eaao6760.

[78] Peter AM Ruijten, Antal Haans, Jaap Ham, and Cees JH Midden. 2019. Perceived human-likeness of social robots: testing the Rasch model as a method for measuring anthropomorphism. *International Journal of Social Robotics* 11 (2019), 477–494.

[79] Ofir Sadka, Alon Jacobi, Andrey Grishko, Udi Lumnitz, Benny Megidish, and Hadas Erel. 2022. " By the way, what's your name?": The Effect of Robotic Bar-stools on Human-human Opening-encounters. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–6.

[80] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1–8.

[81] Nicole Salomons, Kaitlynn Taylor Pineda, Adérónké Adéjàre, and Brian Scassellati. 2022. " We Make a Great Team!": Adults with Low Prior Domain Knowledge Learn more from a Peer Robot than a Tutor Robot. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*. 176–184.

[82] Nicole Salomons, Sarah Strohkorb Sebo, Meiying Qin, and Brian Scassellati. 2021. A minority of one against a majority of robots: Robots cause normative and informational conformity. *ACM Transactions on Human-Robot Interaction (THRI)* 10, 2 (2021), 1–22.

[83] Nicole Salomons, Michael Van Der Linden, Sarah Strokhorb Sebo, and Brian Scassellati. 2018. Humans conform to robots: Disambiguating trust, truth, and conformity. In *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 187–195.

[84] Roger C Schank and Robert P Abelson. 2013. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Psychology Press.

[85] Nicole Shechtman and Leonard M Horowitz. 2003. Media inequality in conversation: how people behave differently when interacting with computers and people. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 281–288.

[86] Muzafer Sherif. 1988. *The robbers cave experiment: Intergroup conflict and cooperation.[Orig. pub. as Intergroup conflict and group relations]*. Wesleyan University Press.

[87] Masahiro Shiomi, Hidenobu Sumioka, and Hiroshi Ishiguro. 2020. Survey of Social touch interaction between humans and robots. *Journal of Robotics and Mechatronics* 32, 1 (2020), 128–135.

[88] Elaine Short, Justin Hart, Michelle Vu, and Brian Scassellati. 2010. No fair!! an interaction with a cheating robot. In *2010 5th acm/ieee international conference on human-robot interaction (hri)*. IEEE, 219–226.

[89] Eliot R Smith, Selma Šabanović, and Marlena R Fraune. 2021. Human-robot interaction through the lens of social psychological theories of intergroup behavior. (2021).

[90] Patric R Spence, David Westerman, Chad Edwards, and Autumn Edwards. 2014. Welcoming our robot overlords: Initial expectations about interaction with a robot. *Communication Research Reports* 31, 3 (2014), 272–280.

[91] Sarah Strohkorb Sebo, Ling Liang Dong, Nicholas Chang, and Brian Scassellati. 2020. Strategies for the inclusion of human members within human-robot teams. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 309–317.

[92] Sarah Strohkorb Sebo, Margaret Traeger, Malte Jung, and Brian Scassellati. 2018. The ripple effects of vulnerability: The effects of a robot's vulnerable behavior on trust in human-robot teams. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 178–186.

[93] Hamish Tennent, Solace Shen, and Malte Jung. 2019. Micbot: a peripheral robotic object to shape conversational dynamics and team performance. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 133–142.

[94] Michael Tomasello. 2014. The ultra-social animal. *European journal of social psychology* 44, 3 (2014), 187–194.

[95] John C Turner, Rupert J Brown, and Henri Tajfel. 1979. Social comparison and group interest in ingroup favouritism. *European journal of social psychology* 9, 2 (1979), 187–204.

[96] Jens Van Lier, Russell Revlin, and Wim De Neys. 2013. Detecting cheaters without thinking: Testing the automaticity of the cheater detection module. *PloS one* 8, 1 (2013), e53827.

[97] Jan Verplaetse, Sven Vanneste, and Johan Braeckman. 2007. You can judge a book by its cover: the sequel.: A kernel of truth in predictive cheating detection. *Evolution and Human Behavior* 28, 4 (2007), 260–271.

[98] Anna-Lisa Vollmer, Robin Read, Dries Trippas, and Tony Belpaeme. 2018. Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Science robotics* 3, 21 (2018), eaat7111.

[99] Auriel Washburn, Akanimoh Adeleye, Thomas An, and Laurel D Riek. 2020. Robot errors in proximate HRI: how functionality framing affects perceived reliability and trust. *ACM Transactions on Human-Robot Interaction (THRI)* 9, 3 (2020), 1–21.

[100] Christian JAM Willemse and Jan BF Van Erp. 2019. Social touch in human–robot interaction: Robot-initiated touches can induce positive responses without extensive prior bonding. *International journal of social robotics* 11, 2 (2019), 285–304.

[101] Robert H Wortham, Andreas Theodorou, and Joanna J Bryson. 2016. What does the robot think? Transparency as a fundamental design requirement for intelligent systems. In *Ijcai-2016 ethics for artificial intelligence workshop*.

[102] Sean Ye, Glen Neville, Mariah Schrum, Matthew Gombolay, Sonia Chernova, and Ayanna Howard. 2019. Human trust after robot mistakes: Study of the effects of different forms of robot communication. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–7.

[103] Xuan Zhao and Bertram F Malle. 2022. Spontaneous perspective taking toward robots: The unique impact of humanlike appearance. *Cognition* 224 (2022), 105076.

[104] Shujie Zhou and Leimin Tian. 2020. Would you help a sad robot? Influence of robots' emotional expressions on human-multi-robot collaboration. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1243–1250.

[105] Jakub Złotowski, Hidenobu Sumioka, Friederike Eyssel, Shuichi Nishio, Christoph Bartneck, and Hiroshi Ishiguro. 2018. Model of dual anthropomorphism: The relationship between the media equation effect and implicit anthropomorphism. *International Journal of Social Robotics* 10 (2018), 701–714.