Telematics and Informatics 78 (2023) 101950

Contents lists available at ScienceDirect

Telematics and Informatics

journal homepage: www.elsevier.com/locate/tele

A mini imitation game: How individuals model social robots via behavioral outcomes and social roles

Kun Xu

College of Journalism and Communications, University of Florida, 3065 Weimer Hall, 1885 Stadium Rd, Gainesville, FL 32611, USA

ARTICLE INFO

Keywords:
Human-robot interaction
Media Equation
The Computers are Social Actors paradigm
Social presence
Social cognitive theory
Modeling effects

ABSTRACT

In past works, social robots have been designed to mimic human appearances and behavior. However, little is known about how human beings may imitate social robots. Drawing on social cognitive theory and the Media Equation, this study focuses on the modeling effects of social robots in an environment protection context. A lab experiment ($N = 128$) with a between-subjects factorial design was conducted to examine how social robots' behavioral outcomes and social roles affected individuals' modeling behavior. This study suggested that social robots' positive behavioral outcomes were effective in evoking users' modeling tendencies serially through social presence and identification or only through identification. Robots' mere presentation of behavior with no outcomes exerted effects serially through social presence and identification. Additionally, assigning social robots an instructor role led to users' modeling behavior serially through users' perception of robots' expertise and credibility. The study analyzed the psychological mechanisms behind users' modeling behavior.

1. Introduction

In his article *Computing Machinery and Intelligence*, Turing (1950) proposed one of the most pivotal thought experiments in the history of human–computer interaction (HCI): the imitation game. The basic idea of the thought experiment was that if a machine assumes the part of a human role and tricks a human interrogator into believing that the machine is indeed a person, then the machine can be considered capable of imitating humans and be perceived as intelligent. Since then, ample work has been done to address the question, "Can machines think?" Searle (1980) proposed the Chinese Room Experiment in his work *Minds, Brains, and Programs*, which suggested that a computer's imitation of human communication cannot be equated with real mastery of human language. Later, Suchman (2007) depicted the significance of situated action in HCI and argued that every meaningful design of interactions between humans and machines must "account for specific local contingent determinants" (p. 84). In other words, machines cannot be considered intelligent if they cannot detect the breaches in communication, which are considered prevalent and trivial in human–human interaction.

Aligned with these philosophical contemplations of whether machines can mimic or replace human intelligence in communication, the dominant paradigm in HCI research, especially in human-robot interaction (HRI), has been focused on setting humans as role models and devising machines to simulate human personalities, emotions, and actions. For example, HuggieBot, developed by the Max Planck Institute for Intelligent Systems at ETH Zürich, was designed to provide soft and warm hugs to those who do not regularly receive hugs from other humans (Fadelli, 2021). Facebook developed a bot that was trained to mimic how people adjust their

E-mail address: kun.xu@ufl.edu.

https://doi.org/10.1016/j.tele.2023.101950
Received 29 October 2021; Received in revised form 22 December 2022; Accepted 27 January 2023
Available online 30 January 2023
0736-5853/

expressions in daily conversations (Reynolds, 2017). Furthermore, prior literature has identified a range of social cues, such as human voices (Cho et al., 2019), human appearances (Barco et al., 2020; Li et al., 2010), humanlike movements (Salem et al., 2013), and anthropomorphic language styles (Hoffmann et al., 2020; Straten et al., 2020) as important elements in developing social robots.

While endowing machines with humanlike behaviors and appearances reflects a human-centric approach, researchers have not fully discussed the implications of the opposite lens: can machines be treated as role models? Just as Groom and Nass (2007) suggested, while developers have attempted to create humanlike robots, they "sometimes overlooked what makes robots special" (p. 494). Considering that social robots afford the programmability of utilizing visual, verbal, and kinesic cues to exert social influence in public spaces (Castellano et al., 2021; Thunberg & Ziemke, 2020), this study explores the contingencies under which humans treat robots as role models and hence imitate robots' pro-social behavior.

Instead of following the previous paradigm that delves into how machines can be designed to mimic human beings, this study draws on the integration of social cognitive theory (SCT) and the Media Equation (TME) to inquire into the potentiality of using social robots to evoke humans' modeling behavior. Specifically, this study investigates an environment protection scenario and examines how social robots may exert modeling effects on individuals' actual behavioral change via the influence of behavioral outcomes and social roles.

Different from previous literature, this line of research could make important theoretical contributions. First, prior research on SCT has been primarily focused on individuals' learning behavior in interpersonal or mediated communication contexts. This study is one of the first to evaluate whether SCT holds its explanatory power and predictive power in HRI. Second, prior SCT research has established the relationship between individuals' expected behavioral outcomes and their modeling behavior (Bandura, 2009), however, little is known about the psychological mechanism underlying this process in HRI contexts. To understand the explanatory power of SCT in HRI, this study parses out individuals' psychological processing of social robots' modeled behavior and evaluates the roles that social presence and identification play in this communication process. Third, research on TME has described how individuals apply social scripts, such as etiquette rules and similarity attraction rules, to HCI (Reeves & Nass, 1996), however, whether individuals apply the social learning script to HRI remains to be explored. Hence, this study tests the extension of TME and seeks to fathom how individuals may learn from robots that present different behavioral outcomes and undertake different social roles even when they are aware that their communication partner is a machine rather than a human.

2. Literature review

2.1. Vicarious learning in social cognitive theory (SCT)

SCT focuses on "psychosocial functioning in terms of triadic reciprocal causation" (Bandura, 2009, p. 94). According to the theory, personal, environmental, and behavioral factors influence each other in this triadic relationship. As part of the personal factors, Bandura (2009) argued that human beings have the capacity for symbolization, self-regulation, self-reflection, and vicarious learning. As this study centers on how users may learn from the modeling behavior of social robots, the psychological mechanism underlying vicarious learning is explained below.

Bandura (2009) suggested that vicarious learning occurs through observing others' behaviors and the consequences of these behaviors. Forming expectations of behavioral outcomes is a key step in this modeling process. Outcome expectation refers to a person's perception of the likely consequences of an action (LaRose et al., 2001). Bandura (2009) argued that outcome expectations function as either incentives or disincentives for behaviors. Positive outcome expectations incentivize individuals to conduct behavior that leads to valued outcomes. Negative outcome expectations inhibit people from engaging in certain behaviors, as these behaviors may bring about undesirable consequences.

Behavioral outcomes have been treated as a central focus in SCT research. Garrett and Ganziger (2008) found that the expected utility of the Internet was a positive predictor of personal Internet use, and the expected restrictions on computer use were negative predictors of personal Internet use during work. The effects of outcome expectations have also been found in research on media and body image. Hendricks (2002) found that the more women compared themselves with the characters on TV, the more they endorsed a thin body image. Overall, based on the tenets of SCT, individuals are more likely to present modeled behavior if the behavior leads to desired outcomes, but less likely to carry out the behavior if the behavior generates unrewarding outcomes (Bandura, 2009).

While plenty of evidence has clarified the effects of behavioral outcomes on people's modeling behavior via mass media channels (Nabi & Clark, 2008; Zimmerman et al., 2012), limited research has investigated whether these modeling effects can be transferred to HRI (Zanatto et al., 2020). Considering that individuals' communication partners switch from media characters to social robots in this process, TME can be incorporated to better tackle the question.

2.2. The media Equation (TME)

TME was derived from the Computers are Social Actors (CASA) paradigm, in which Nass and colleagues found that humans apply the social scripts of interpersonal communication to HCI. These social scripts include, but are not limited to, the politeness rule, gender stereotypes, and the personality attraction rule (see Nass & Moon, 2000). Drawing on these findings, Reeves and Nass (1996) described that individuals' responses to media technologies are fundamentally social and natural. As "new media engage old brains" (p. 68), media users equate interaction with media technologies to interaction with humans, especially when these technologies demonstrate social cues, such as human voices and language.

TME has been applied to a wide range of technologies, including televisions, computer agents, chatbots, and even smartphones (Carolus et al., 2019; Edwards et al., 2019; Horstmann et al., 2018; Westerman et al., 2019). Here, TME is used to explain users' social reactions to social robots. Social robots have been conceptualized through various angles. For example, Duffy (2003) highlighted the physical embodiment nature and defined a social robot as “a physical entity embodied in a complex, dynamic, and social environment sufficiently empowered to behave in a manner conducive to its own goals and those of its community” (p. 177). Zhao (2006) provided a broader explication and defined humanoid social robots as “human-made autonomous entities that interact with humans in a humanlike way” (p. 405). Similarly, Fox and Gambino (2021) referred to humanoid social robots as “human-made technologies that can take physical or digital form, resemble people in form or behavior to some degree, and are designed to communicate with people” (p. 295). Although scholars diverge on their explications, social robots at least feature a certain extent of automation and the capacity of engaging in social interactions with their users (Lee et al., 2006; Xu, 2019).

Research on TME has examined users’ different types of social responses, including users’ trust in technologies, perceived attraction of technologies, and conformity tendencies. Focusing on behavioral intentions, Xu and Lombard (2017) found that even when users were aware that they were interacting with a group of computer agents, they reported strong conformity intentions when these computer agents imposed uniform group norms. Revolving around actual behavioral change, Hoffman and Krämer (2021) found that compared to those who were not touched by a social robot, participants who were touched complied more with the robot’s request. Horstmann et al. (2018) found that when receiving a social robot’s objection, participants treated it as a social actor and waited longer to switch the robot off.

Existing knowledge about vicarious learning and social responses may provide a fruitful starting point for investigating the modeling effects of social robots. Based on SCT and TME, it is reasonable to postulate that when a social robot demonstrates social cues, such as gestures, movements, voices, and eye contact, users may develop social responses to the robot. Such social responses may further allow individuals to transfer social learning behavior to HRI. Considering that Karutz and Bailenson (2015) noted that social learning may take effect not only in the form of human models but also symbolic models such as digital characters, we postulate that individuals may model social robots’ behavior based on their perceived behavioral influence (i.e., expected outcomes of robots’ behavior). Furthermore, as prior meta-analysis has suggested that individuals’ behavioral intentions have small-to-medium-sized effects on actual behavioral change in social psychology (Webb & Sheeran, 2006), we propose the following hypothesis to test the indirect effects of social robots’ behavioral outcomes on individuals’ modeling behavior:

> ***H1***: Compared to negative behavioral outcomes, a social robot that demonstrates positive behavioral outcomes will be more likely to lead to individuals’ modeling behavior through their perceived behavioral influence and modeling intention.

2.2.1. Perception-behavior link

Past works have documented how modeling behavior may occur without reinforcing the succeeding consequences. (Bandura et al., 1961). The original Bobo Doll experiment demonstrated that, even without rewards or punishments, an adult’s aggressive behavior still led to children’s violent behavior (Bandura et al., 1961), implying that individuals’ mere observation of the behavior could lead to their imitation. This perspective was further supported in research on perception-behavior link (Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001), where scholars have suggested that perceiving others’ behavior automatically creates the tendency to mimic the behavior. Chartrand and Bargh (1999) found in their experiment that even minimal interaction with a stranger led to a certain level of imitation, which included mimicking facial expressions, postures, and speech patterns. A possible explanation for the perception-behavior link lies in the mechanism of natural selection: people imitate others’ behavior to reduce potential risks and to seek a sense of belongingness (Chartrand & Bargh, 1999). While the perception-behavior link is interpreted to be wired into our brain, research has noted that disincentives can generate counter forces on imitation and override the link (Bargh & Fergusson, 2000; Macrae & Johnston, 1998). That is, compared to negative behavioral outcomes, mere observation of a robot’s behavior should be more likely to lead individuals to model the behavior. Thus, based on SCT, TME, and the perception-behavior link, we propose the following hypothesis to test the indirect effects of social robots’ mere behavioral presentation on individuals’ modeling behavior.

> ***H2***: Compared to negative behavioral outcomes, a social robot that demonstrates behavior without presenting behavioral outcomes will be more likely to lead to individuals’ modeling behavior through their perceived behavioral influence and modeling intention.

2.3. Social presence and identification

Research on the psychological mechanism of vicarious learning suggests that it is indeed the identification with the model that facilitates imitation. Identification is defined as the degree to which individuals relate to the model and feel similar to it (Fox & Bailenson, 2009). The perceived similarities can be the model’s gender, personality, social roles, or skills (Karutz & Bailenson, 2015). For example, prior research has suggested that viewers were more likely to identify with television models that were of the same gender as them (Andsager et al., 2006). Fox and Bailenson (2009) advanced the Proteus Effect and noted that seeing a virtual representation of self being rewarded for losing weight encouraged individuals to exercise, while observing virtual representations of others was not sufficient to trigger more exercise, which confirmed the significance of identification with the model even in virtual environments.

The role of identification in the modeling effects was further noted in Nabi and Clark’s (2008) study, where they found that favored TV characters may infuse their behaviors with positive value, despite the risks of the behaviors. If the effects of favored TV characters outweigh the negative outcome expectations, then exposing viewers to risky behaviors would not be sufficient to counterbalance the impact of identification. Therefore, it can be implied that identification plays a mediating role between the perceived behavioral outcomes of the model and users' modeling tendencies.

Despite the importance of identification in individuals' vicarious learning from other human models, given that humans and machines have ontological divides (Edwards, 2018; Guzman, 2020) and do not share gender or personalities in nature, it is assumed that users' identification with social robots would not automatically arise unless users perceive social robots as if they were real people in the first place. In other words, to develop identification with social robots, the stepping-off point is that users need to perceive robots as social beings.

It is here where the concept of social presence may bridge the relationships among perceived behavioral outcomes, identification, and users' modeling behavior. Social presence has been defined as "a psychological state in which virtual (para-authentic or artificial) social actors are experienced as actual social actors in either sensory or non-sensory ways" (Lee, 2004, p. 45). A community of presence scholars at the 2000 Presence Conference conceptualized social presence as the experience where a person partially or fully overlooks the mediating nature of technology and feels as if they were communicating with other people (ISPR, 2000). Lombard and Ditton (1997) further distinguished two types of social presence: social-actor-within-medium presence and medium-as-social actor presence. While the former type often emerges in mediated communication contexts where users respond to media characters or avatars as real social actors, the latter type occurs more often in HCI contexts where users directly respond to the cues presented by the media technologies per se. Based on the definitions, perceiving social robots as if they were social beings has been considered a typical example of medium-as-social-actor presence, as users partially overlook the nature of machines and respond to them as real people (Lombard & Xu, 2021; Xu & Liao, 2020).

Social presence has been found to be a mediator between social cues and users' social responses to social robots. Jung and Lee (2004) found that the physical embodiment of the zoomorphic robot AIBO had indirect effects on users' positive evaluations of the robot through social presence. Lee et al. (2006) found that when participants interacted with the AIBO that exhibited complementary personalities to their own, they experienced stronger social presence than when they interacted with a robot that exhibited the same personalities as theirs. Stronger social presence further led to the perceived intelligence and social attraction of the robot.

While research has established the links between robots' behavioral cues and social presence (Bevan & Fraser, 2015) and the links between social presence and social responses (Lee et al., 2005; Lee et al., 2006), to our knowledge, no prior literature has elucidated the psychological mechanism of how individuals model social robots' behavior through social presence and identification. In this study, it is postulated that the relationship between social presence and identification may help illuminate social robots' modeling effects, as identification plays a mediating role in vicarious learning and social presence is a necessary precursor to individuals' identification with social robots. Given that individuals' actual behavior change may occur due to their behavioral intentions (Ajzen, 1991; Webb & Sheeran, 2006), the following hypothesis is proposed to test the psychological processing of social robots' modeled behavior.

> ***H3***: Compared to negative behavioral outcomes, a social robot's positive behavioral outcomes will be more likely to lead to individuals' modeling behavior serially through their perceived behavioral influence, social presence, identification, and modeling intention.
>
> ***H4***: Compared to negative behavioral outcomes, a social robot that demonstrates behavior without presenting behavioral outcomes will be more likely to lead to participants' modeling behavior serially through social presence, identification, and their modeling intention.

2.4. *Social roles of social robots*

One question that has continuously been discussed among HRI scholars concerns the social roles robots should assume in public sphere. Back in the 1990s, Nass et al. (1995) examined the degree to which individuals were willing to accept computers in human roles. They found that users were more likely to accept computers in routinized roles (e.g., bank tellers, accountants) than in interpretive roles (e.g., newspaper reporters, editorial writers) or personal roles (e.g., babysitters, judges). Since then, social robots have been tested in various positions including museum guides, information kiosk assistants, and food delivery staff (Mutlu & Forlizzi, 2008; Shiomi et al., 2006). While most of these positions are routinized social roles, research has advanced to a level where robots are assigned more interpretive and personal roles. Lee et al. (2017) found that social robots can be designed as healthcare advisors to effectively augment patients' compliance. Looije et al. (2010) found that social robots can persuade older adults to conform to doctors' suggestions by showing empathy and demonstrating emotions. As social robots become more personal and intelligent, their social roles begin to "fall into a hybrid category between strictly utilitarian and affective" (Shaw-Garlock, 2009, p. 250).

Among the various social roles assigned to social robots, researchers have debated whether social robots should undertake a fellow role with heightened peer relations with users, or a specialist role with authority and area expertise, as both approaches have been found effective in evoking users' positive attitudes toward robots. The former line of research centers on how users may form equal relations with social robots and maintain an in-group relationship with them (Eyssel & Kuchenbrandt, 2012). The latter line of research has suggested that social robots can serve as instructors with knowledge and expertise to promote attitudinal and behavioral change. Examples include robots playing a coaching role in encouraging adults to undertake physical activities (Čaić et al., 2020) or a teacher role in enhancing students' learning performances (Edwards et al., 2016). In this process, social roles with specialization evoke users' social responses through perceived expertise. As Liew and Tan (2018) pointed out, a specialist virtual agent could augment online users' purchase intentions through users' interpretations of the agent expertise. Based on the intricacy of the different effects produced by robots as instructors and robots as fellows in prior research, this study continues to explore this question and seeks to understand how social robots in different social positions affect users' modeling tendencies. We propose the following hypothesis.

H5: Compared to a fellow role, a social robot in an instructor's role will be more likely to lead to individuals' modeling behavior through their perceived expertise and modeling intention.

Individuals' psychological processing of robots' social roles is further contingent upon their trust in technologies. Past research has suggested that when a robot manifested expertise, participants developed more affective trust in it than when expertise was not emphasized (Sah et al., 2011). As establishing trust in social robots further facilitates users' intention of behavioral change (You & Robert, 2018), it can be postulated that trust is a key component in explaining how users' perception of robots' expertise influences their modeling intention. Based on above-mentioned literature, we further propose the following hypothesis.

H6: Compared to a fellow role, a social robot in an instructor's role will be more likely to lead to individuals' modeling behavior serially through their perceived expertise, trust, and modeling intention.

Combining the literature on vicarious learning and social roles, this study further investigates the interaction between a social robot's behavioral outcomes and its social roles. It is speculated that a social robot presenting positive behavioral outcomes may exert strong effects on users' modeling behavior when the robot is assigned a credible and persuasive social role. However, as it remains to be explored whether it is the instructor role or the fellow role that creates greater modeling tendencies, we propose the following research question.

RQ1: How will a social robot's behavioral outcomes interact with its social roles in affecting participants' 1) modeling intention and 2) modeling behavior?

3. Method

3.1. Participants

Based on a priori estimation with $\beta = 0.2$, effect size $f = 0.03$, and $a = 0.05$, power analysis suggested that at least 111 participants were needed for the sample size. A total of 138 participants were recruited from a large public university in the US to participate in a lab experiment[1]. They were informed that researchers were seeking to test the usability of a prototype social robot and a new garbage recycling system on campus. The garbage recycling tasks were used, as prior research had tested the feasibility of using social robots for promoting children's attitudes toward recycling (Castellano et al., 2021). Participants were recruited either through the SONA system and offered extra credit or through advertisements posted on campus boards and rewarded with $10 gift cards. Of the 138 cases, eight were excluded due to technical problems in the experiment (e.g., the robot failed to recognize human faces due to masks, or the robot froze during the experiment). Two cases were removed as univariate or multivariate outliers. Thus, 128 participants were included in the final analyses. Their average age was 20.08 years ($SD = 1.51$). Among them, 31 were males (24.2 %), 96 were females (75.0 %), and one selected "other" (0.7 %).

3.2. Experimental stimulus

The social robot *NAO*, created by SoftBank Robotics, was used in the experiment. *NAO* has 25 degrees of freedom, which enables it to move its head, limbs, hands, and body. It is also equipped with seven touch sensors, two cameras, and four speakers to interact with the environment. The software Choregraphe was used to program its messages, behavior, visual recognition tasks, and reactions. Its own synthetic voice was used across all conditions. *NAO* has been applied in university settings in past HRI works (Hoffman & Krämer, 2021; Li et al., 2017).

A 3 (positive outcomes, negative outcomes, or no outcomes) X 2 (instructor, fellow) between-subjects factorial design was used. Participants were randomly assigned to one of the conditions. To manipulate the robot's behavioral outcomes, the robot was programmed to present the same behavior, but with positive, negative, or no verbal messages. Verbal messages were used as Bandura (1977) suggested models can be established through verbal instructions, such as explaining the outcomes of a group of actions.

In the positive outcome conditions, the robot stated the positive consequences of its recycling behavior. For example, after the robot recycled plush toys into the container for recyclable waste, the robot delivered the message, "As producing plush toys consumes an extraordinary level of electricity and water, sorting plush toys into recyclable waste will bring about positive influence. It can conserve energy and help the people in developing countries." In the negative outcome conditions, the robot informed participants of the negative consequences of its behavior, "Actually sorting plush toys into recyclable waste does not ensure that it can reproduce high-quality products. It can bring about negative influence. Bleaching the plush toys or clothing will expose workers to harsh working conditions that may deteriorate their health." A total of four messages related to plush toys, green plants, plastic bags, and hair dye bottles were provided to participants during the experiment (see Appendix A of the Supplementary Materials for all the messages: https://osf.io/bhsr6/?view_only=e148010832b74c3d9246ffa45a1e7432). In the no outcome conditions, the robot did not provide

[1] Given the influence of the COVID-19 pandemic and reviewers' constructive feedback, the author withdrew the submission in early 2022, re-ran the experiment from January to May in 2022, and submitted the manuscript in December 2022. Participants wore masks and followed the university COVID-19 policies during the experiment. Lab assistants invested additional time disinfecting the robot, lab computers, desks, and chairs between each session of the experiment.

any verbal messages after its recycling behavior.

To manipulate the robot's social roles, in the instructor conditions, the robot introduced itself with the following message, "I was invited by the university to help build a new garbage recycle system. I work as an instructor with expertise on garbage recycle here. I have been trained in different organizations to master the garbage recycle skills, and to teach the necessary knowledge. I am here to advise on the new garbage recycle system." In the fellow conditions, the message was, "I was brought here as a fellow to learn about a new garbage recycle system. I did not receive any formal training, and I only had limited knowledge about garbage recycle skills. I am here to help test the new garbage recycle system".

3.3. Research design and procedures

After participants entered the lab and signed the consent forms on a desktop computer, they were first asked about some demographic information and prior garbage recycling knowledge via the software Qualtrics. Participants were then asked to work on a garbage recycling task.

In the task, participants were asked to sort green plants, plush toys, plastic bags, and hair dye bottles into four recycling categories: compost waste, hazardous waste, recyclable waste, and residual waste. The recycle categories were sampled from the official trash sorting guidelines released by the Kyoto, Berlin, and Shanghai municipal governments for stimulus sampling purpose (Reeves et al., 2016). Meanwhile, these categories were used to minimize the effects of participants' existing knowledge about campus waste recycling principles.

After participants made their initial choices, they were led to a table where the social robot *NAO* was sitting. Four objects that represented each of the four garbage items, along with four empty recycling bins, were placed in front of the robot (see Fig. 1). Participants stood at a position that was about 25 to 30 inches away from the robot. They were given a script detailing the steps they needed to follow and the basic interaction techniques they could use to communicate with the robot. At the beginning, every participant was asked to greet the robot (e.g., participants said, "how are you," "nice to meet you," or "how is it going" to the robot). The robot responded with "fine," "I'm good," or "super!" This step was designed to familiarize participants with the robot's reaction speed and voice features.

After the greetings, participants were instructed to ask the robot why it was brought to the university. The robot responded to participants with the messages about its social roles (i.e., instructor vs fellow). Next, participants asked the robot to conduct the garbage recycling task. The robot was programmed to introduce what it could do first and then ask participants to choose a garbage item they would like it to sort. After participants told the robot the item they selected, the robot confirmed participants' choice and let participants show the image of the item to it. When the robot recognized the item via its visual recognition system, it told participants to pick the object that represented the item and put the object in its hand. Meanwhile, the robot stretched out one of its arms and opened its hand. After participants placed the item into the robot's hand, the robot grasped the object, moved its arms, and threw the

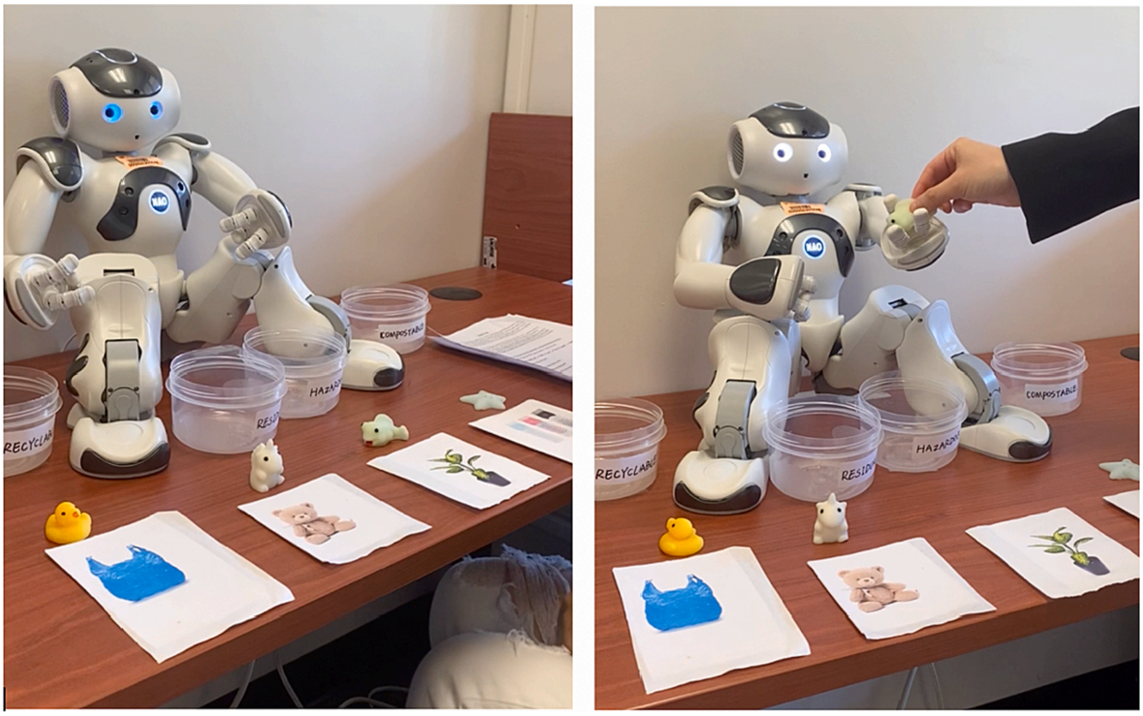

Fig. 1. Experimental setting.

item into the matching recycling bin. The robot then either delivered the messages about the outcomes of its behavior (i.e., positive outcomes vs negative outcomes), or returned to its initial position without any follow-up message (i.e., no outcome condition).

After the robot sorted the first garbage item, participants were asked to let the robot sort the next item based on their selection. Participants repeated this process until the robot finished sorting all four items (see Fig. 2). After the robot finished the whole process, participants were led back to the computer and asked to work on the recycling task again to make their final decisions. They were informed that the robot's recycling behavior was not guaranteed to be correct. Finally, participants filled out the rest of the questionnaire and received the debriefing form indicating the true experiment purposes. It should be noticed that the experiment was not designed to test whether participants would recycle trash, but whether and how participants would imitate the way the robot recycled trash.

3.4. Measures

The measure of social presence ($M = 5.29$, $SD = 1.01$, $\alpha = 0.86$) was adapted from previous measures of social presence in the contexts of HRI (Lee et al., 2006). Participants were asked to report on a seven-point Likert-type scale with six items (1 = not all, 7 = very much). Examples of the items included "how much did you feel as if you were interacting with an intelligent being?" and "how much did you feel as if you were together with an intelligent being?".

The measure of identification ($M = 5.64$, $SD = 0.74$, $\alpha = 0.70$) was adapted from the previous measure of identification (Cohen, 2001). Participants were asked to report on a seven-point Likert-type scale with seven items (1 = strongly disagree, 7 = strongly agree). Examples of the items included "I tend to understand the reasons why the robot did what it did" and "I think I had a good understanding of the robot".

The measure of trust ($M = 5.78$, $SD = 0.89$, $\alpha = 0.79$) was adapted from previous measures of trust in HCI (Gong, 2008; Wheeless & Grotz, 1977). Participants were asked to report on five seven-point semantic differential scale items. Examples of the items included "untrustworthy – trustworthy" and "unreliable – reliable".

The measure of modeling intention ($M = 5.76$, $SD = 1.00$, $\alpha = 0.82$) was adapted from previous measures of conformity intention and openness to influence (Kim & Park, 2011; Nass et al., 1996). Participants were asked to report on a seven-point Likert-type scale with five items (1 = strongly disagree, 7 = strongly agree). Examples of the items included "I was open to influence from the robot" and "I desired to reach agreement with the robot".

Modeling behavior ($M = 6.93$, $SD = 4.54$) was operationalized as the change of participants' choices to align with the robot's garbage recycling behavior. It was adapted from previous measures of conformity behavior (Lee & Nass, 2002; Takayama et al., 2009) and measured by calculating the percentage of the number of changes that participants made to match the robot's behavior and the total number of the inconsistencies between participants' initial choices and the robot's recycling behavior. For example, if a participant's initial recycling decisions were completely inconsistent with the robot's behavior (i.e., all four recycling actions were inconsistent with the robot's behavior) and after observation, the participant changed one recycling decision to align with the robot's behavior, then the modeling behavior was coded as 25 %. If participants did not make any change or did the opposite of what the robot presented, their score was coded as zero, meaning that no modeling behavior occurred in this process. The percentage was converted to a 10-point system.

Perceived behavioral influence ($M = 4.45$, $SD = 0.69$) was measured by asking participants how much positive or negative influence the robot's garbage recycling behavior could lead to. Participants reported on a five-point semantic-differential scale item (1 = negative, 5 = positive).

Perceived expertise ($M = 4.26$, $SD = 0.85$, $\alpha = 0.87$) was measured by adapting a previous measure of perceived expertise in HCI (Koh & Sundar, 2010). Participants were asked to report on a five-point Likert-type scale with three items (1 = strongly disagree, 5 = strongly agree). Examples of the items included "the robot is an expert on garbage recycle" and "the robot specializes in garbage recycle".

In addition to demographic information, participants' attitudes toward social robots and self-efficacy were used as control variables. The measure of self-efficacy ($M = 6.16$, $SD = 0.72$, $\alpha = 0.79$) was adapted from the previous measure of self-efficacy (Compeau & Higgins, 1995). Participants were asked to respond to five seven-point Likert-type items. A sample item was "I am confident that I could complete the job of trash sorting if someone else helps me get started." Participants' attitudes toward social robots ($M = 1.95$, $SD =$

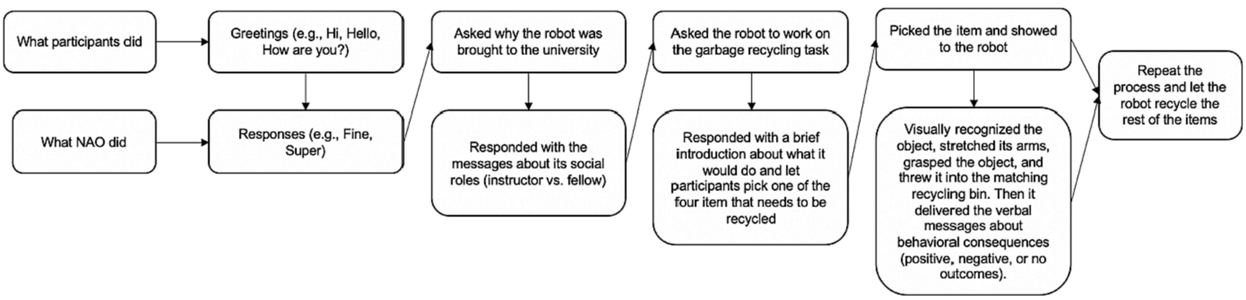

Fig. 2. Experimental procedures.

0.80, $a = 0.78$) were measured by adapting the previous measure (Nomura et al., 2006). Participants reported on five five-point Likert-type scale items (1 = strongly disagree, 5 = strongly agree). A sample item was "I would feel uneasy if I was given a job where I had to use robots." (see Appendix B in Supplementary Materials for more information about control variables https://osf.io/bhsr6/?view_only=e148010832b74c3d9246ffa45a1e7432).

3.5. Data analyses

SPSS was used to test and examine the hypotheses and research questions. Univariate and multivariate outliers were examined. Multicollinearity was checked. The variables of perceived expertise, perceived behavioral influence, and modeling intention were power transformed due to negative skewness. To test *H1-H6*, the Process macro (Hayes, 2013) in SPSS was used. As behavioral outcomes involve three levels, Hayes and Preacher's (2014) approach to analyzing multi-categorical independent variables in mediation analyses was adopted. Based on Tao and Bucy's (2007) experimental paradigm of integrating effect-labeled media attributes and effect-based psychological states, the mediation analyses included manipulation check variables (i.e., perceived behavioral influence, perceived expertise) as necessary mediators. Thus, model 6 was used to run serial mediation. To answer *RQ1*, two-way ANCOVA was used.

4. Results

Manipulation checks were conducted. Results from T-tests suggested that participants in the positive outcomes conditions ($M = 4.60$, $SD = 0.50$) perceived the robot's behavior to have significantly more positive influence than those in the negative outcomes conditions ($M = 4.26$, $SD = 0.73$), $t(81) = 2.41$, $p = .018$, Cohen's $d = 0.54$. Additionally, participants perceived the robot in an instructor role ($M = 4.45$, $SD = 0.57$) as more knowledgeable and specialized than the one in a fellow role ($M = 4.07$, $SD = 1.01$), $t(119.99) = 2.19$, $p = .031$, Cohen's $d = 0.47$. Thus, the manipulation was successful.

To test *H1*, results from serial mediation suggested that compared to negative behavioral outcomes, positive behavioral outcomes were more likely to have indirect effects on participants' modeling behavior through perceived behavioral influence and modeling intention, $B = 0.18$, $LLCI = 0.02$, $ULCI = 0.50$. Thus, *H1* was supported.

To test *H2*, results from serial mediation suggested that compared to negative behavioral outcomes, a social robot's presentation of behavior with no outcomes was more likely to have indirect effects on participants' modeling behavior through perceived behavioral influence and modeling intention, $B = 0.17$, $LLCI = 0.002$, $ULCI = 0.43$. Thus, *H2* was supported.

To test *H3*, results from serial mediation suggested that compared to negative behavioral outcomes, positive behavioral outcomes were more likely to have indirect effects on participants' modeling behavior serially through perceived behavioral influence, identification, and modeling intention, $B = 0.04$, $LLCI = 0.0002$, $ULCI = 0.13$, or serially through perceived behavioral influence, social presence, identification, and modeling intention, $B = 0.02$, $LLCI = 0.001$, $ULCI = 0.06$. Thus, *H3* was supported.

To test *H4*, results from serial mediation suggested that compared to negative behavioral outcomes, the robot's behavior with no outcomes was more likely to have indirect effects on participants' modeling behavior serially through perceived behavioral influence,

Table 1
Results of the Hypotheses.

	Relationship Tested	*B (SE)*	95 % Confidence Interval *LLCI*	*ULCI*	Result
H1	Positive outcomes → perceived behavioral influence → modeling intention → modeling behavior	0.18 (0.13)	0.02	0.50	√
H2	Behavior without outcomes → perceived behavioral influence → modeling intention → modeling behavior	0.17 (0.11)	0.002	0.43	√
H3	Positive outcomes → perceived behavioral influence → social presence → identification → modeling intention → modeling behavior	0.02 (0.02)	0.0012	0.06	√
	Additional results from Process Macro: Positive outcomes → perceived behavioral influence → identification → modeling intention → modeling behavior	0.04 (0.04)	0.0002	0.13	
H4	Behavior without outcomes → perceived behavioral influence → social presence → identification → modeling intention → modeling behavior	0.02 (0.01)	0.0002	0.06	√
H5	Instructor role → perceived expertise → modeling intention → modeling behavior	0.24 (0.14)	0.02	0.60	√
H6	Instructor role → perceived expertise → trust → modeling intention → modeling behavior	0.05 (0.03)	0.002	0.13	√

Note: *LLCI:* lower-level confidence interval; *ULCI*: upper-level confidence interval; *B*: unstandardized coefficient; *SE*: standard error. Modeling intention, perceived behavioral influence, and perceived expertise were transformed for normal distribution. Effects of self-efficacy and attitudes toward interaction with robots were controlled.

social presence, identification, and modeling intention, $B = 0.02$, $LLCI = 0.0002$, $ULCI = 0.06$. Thus, *H4* was supported.

To test *H5*, results from serial mediation suggested that compared to a fellow role, the social robot in an instructor's role was more likely to have indirect effects on participants' modeling behavior through their perceived expertise of the robot and modeling intention, $B = 0.24$, $LLCI = 0.02$, $ULCI = 0.60$. *H5* was supported.

To test *H6*, results suggested that compared to a fellow role, the robot in an instructor role was more likely to have indirect effects on participants' modeling behavior serially through perceived expertise, trust, and modeling intention, $B = 0.05$, $LLCI = 0.002$, $ULCI = 0.13$. *H6* was supported. The results of the tested relationships were shown in Table 1. Other indirect, direct, and total effects tested in each mediation model were shown in Appendix C of the Supplementary Materials https://osf.io/bhsr6/?view_only=e148010832b74c3d9246ffa45a1e7432.

To examine *RQ1*, results from two-way ANCOVA suggested that there was no interaction between the behavioral outcomes and social roles in affecting participants' modeling intention, $F(2, 120) = 0.46$, $p = .63$, or modeling behavior, $F(2, 120) = 1.41$, $p = .25$. However, based on reviewed literature, it should be users' perceptions of the robot's behavioral influence (i.e., expected outcomes) and expertise that exert effects on modeling intention and behavior. Hence, post hoc analyses were conducted to examine the interaction between participants' perceived behavioral influence and perceived expertise. Results from the Process macro Model 1 suggested that although the overall interaction effect on modeling intention was not significant, $B = -0.01$, $p = .13$, $LLCI = -0.02$, $ULCI = 0.002$, the Johnson-Neyman probe suggested that the interaction between perceived behavioral influence and perceived expertise transitioned between statistically significant and non-significant at the value of perceived expertise being 4.46, meaning that for those who did not feel or only moderately felt that the robot was specialized in garbage recycling, the relationship between perceived behavioral influence and participants' modeling intention was statistically significant, while for those who strongly felt that the robot was an expert in garbage recycling, the relationship between perceived behavioral influence and participants' modeling intention was not statistically significant (see Fig. 3). The overall interaction effect on modeling behavior was not statistically significant, $B = -0.0004$, $p = .17$, $LLCI = -0.001$, $ULCI = 0.0002$. The Johnson-Neyman probe suggested that there was no statistical significance transition point within the observed range of the data. The overall psychological mechanisms behind individuals' modeling behavior were presented in Fig. 4.

5. Discussion

While ample research has centered on how social robots can be designed with humanlike appearances, voices, or language styles to mimic human beings and to enhance user experience and communication effectiveness, this study takes an opposite approach and focuses on how social robots can be used as role models to encourage individuals' modeling behavior. Drawing on the tenets of SCT and TME, this study suggests that social robots' presentation of behavior with positive consequences, or mere behavior with no outcomes, exerts indirect effects on individuals' modeling behavior. Such modeling behavior occurs serially through users' perceptions of behavioral outcomes, social presence experience, and their identification with social robots. Moreover, this study follows past calls for more discussion about the effects of robots' social roles and suggests that social robots in an instructor role are more likely than those in a fellow role to evoke users' modeling behavior sequentially through users' interpretation of robots' expertise, trust in social robots, and modeling intention.

5.1. Behavioral outcomes

This study first indicated that compared to negative behavioral outcomes, a social robot's positive behavioral outcomes were more likely to lead to users' modeling behavior serially through their perceptions of behavioral consequences and modeling intention. This

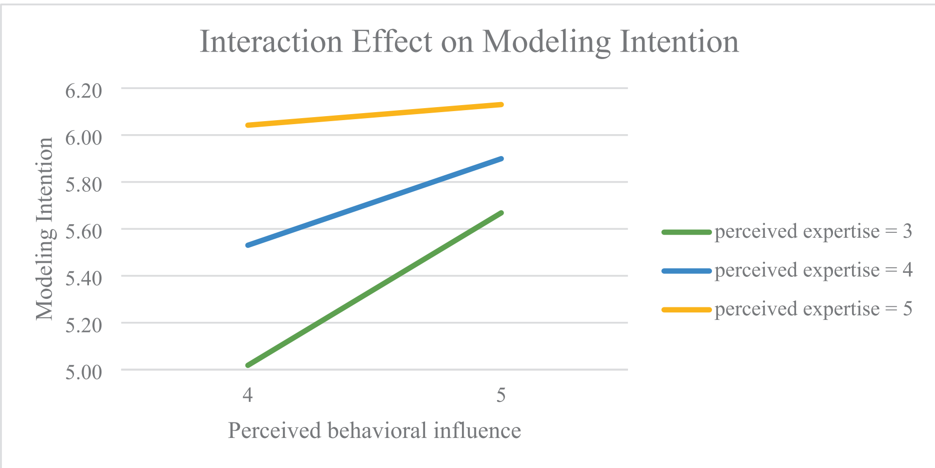

Fig. 3. Interaction between perceived behavioral influence and perceived expertise on modeling intention.

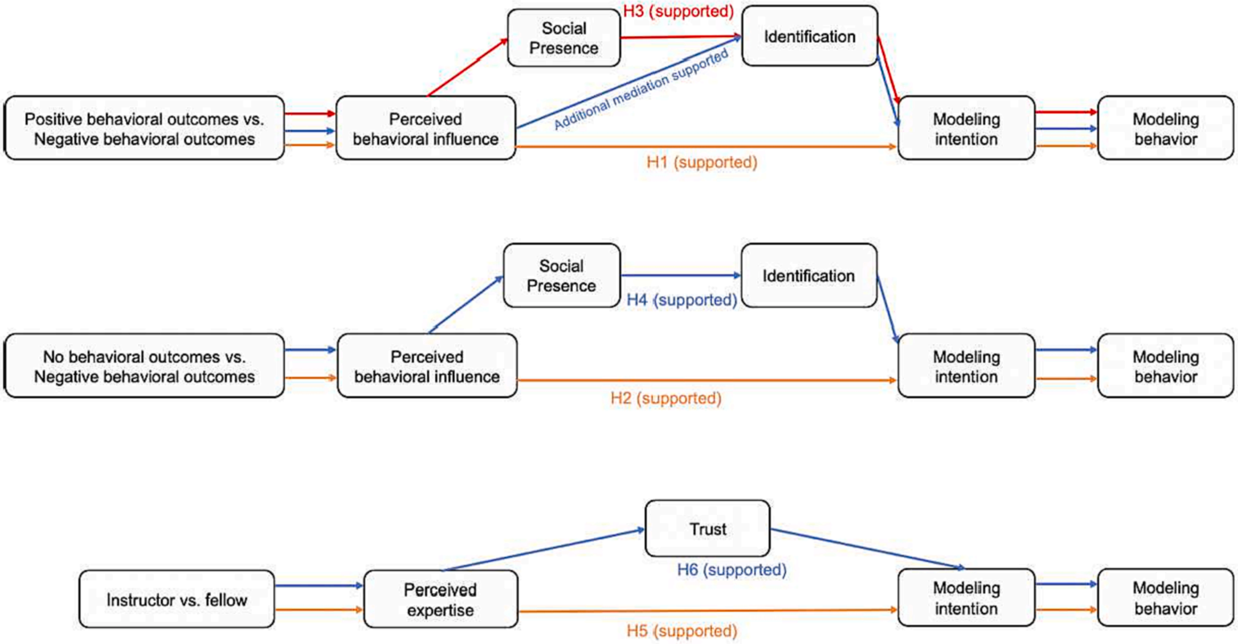

Fig. 4. Findings of the Serial Mediation Analyses.

finding not only supports SCT, where individuals are motivated to model behavior through their expectations of valued consequences rather than undesired ones (Bandura, 2009), but also confirms prior findings that users can act in conformity with digital agents' behavior (Takayama et al., 2009; Zanatto et al., 2020). Given that research on TME has not investigated how individuals may transfer social learning behavior from interpersonal or mass communication contexts to HCI contexts, this finding can contribute to TME in that users are capable of learning behavior by treating social robots as role models just as they do with human models.

The study further indicated that even without behavioral outcomes, merely letting a social robot present the recycling behavior had indirect effects on users' modeling behavior through perceived behavioral consequences and modeling intention. Consistent with the early Bobo Doll experiment finding that without rewards or punishments, children can still model adults' actions (Bandura et al., 1961), the current finding implied that without behavioral outcomes, participants might have assumed that the robot's behavior was correct and thus would lead to positive consequences. Indeed, the mean value of participants' perceived behavioral influence in the behavior with no outcome condition ($M = 4.51$, $SD = 0.76$) was only slightly lower than that in the positive behavioral outcomes condition ($M = 4.60$, $SD = 0.50$), which lent support to the surmise that participants formed positive expectations of *NAO*'s actions. On top of that, the effects of behavior with no outcomes on users' modeling behavior are congruent with the perception-behavior link, where scholars have found that merely perceiving others' actions per se is sufficient to lead to the imitation of those actions (Chartrand & Bargh, 1999).

To better understand individuals' psychological processing of social robots' behavior and the concomitant behavioral outcomes, this study tested the roles of social presence and identification in social robots' modeling effects. Results suggested that when participants perceived positive behavioral consequences, modeling effects occurred either sequentially through users' social presence and identification with the robot, or only through their identification with the robot. That is, aligned with Fox and Bailenson's (2009) research on the role of identification in the Proteus effect, this study indicates that when learning from social robots, the degree to which individuals relate to those technologies is a key factor in determining their modeling tendencies.

Our study further revealed that when a social robot demonstrated behavior with positive outcomes or no outcomes, modeling effects occurred serially through social presence and identification, meaning that perceiving social robots as if they were social actors was an important antecedent for individuals to identify with the social robot. This finding first corroborates previous TME literature in which users' social responses to technologies emerge through their perceptions of technologies as social actors (Lee et al., 2006; Lombard & Xu, 2021). More importantly, this is a novel finding about the application of SCT in HRI contexts because prior SCT research has primarily expounded the key role of identification in vicarious learning, whereas the role of social presence has been largely overlooked. The incorporation of social presence in this study demonstrates that individuals' imitation behavior may hinge upon their experience of artificial actors as real social actors (Lee, 2004).

What merits attention here is that although social presence was not a necessary factor in leading to users' modeling tendencies when the social robot presented behavior with positive outcomes, it played an indispensable role in users' modeling process when the robot presented behavior with no outcomes. This means when a machine explicitly presents its behavioral consequences, users can learn through these consequences without having to perceive it as a social entity. However, when a machine only performs certain actions without demonstrating any succeeding effects, users will be reluctant to model its behavior unless they perceive the machine as an intelligent and communicative social actor. This finding provides a more nuanced understanding of the application of SCT in HRI and reflects that humans' vicarious learning in HRI is similar to that in interpersonal or mass communication contexts when behavioral outcomes are available and perceptible. However, such learning process requires the additional but imperative trigger of social presence when behavioral outcomes are not observable to users.

5.2. *Social roles and interaction effects*

This study further suggested that assigning a social robot an instructor role was more effective than assigning it a fellow role in evoking users' modeling tendencies. This process occurred through the degree to which users perceived the robot to be knowledgeable and specialized in the collaboration task. Although past works have examined the influence of social robots as teachers on students' learning performances (Edwards et al., 2016), this study parses out such influence and highlights that it is indeed users' perceived expertise that triggers their modeling tendencies.

To further unfold the psychological mechanism underlying the relationship between social robots' social roles and users' modeling behavior, this study demonstrated that a social robot's instructor role imposed indirect effects on users' modeling tendencies serially through users' perception of expertise and users' trust in the robot. The finding can add to the current academic conversation about whether a more hierarchical human-robot relationship or a more equal relationship is more effective in promoting pro-social behavior. Deviating from prior research where users relied more on social robots as peers than as supervisors (Hinds et al., 2004), the current finding hints that, at least when a human-robot collaboration task requires certain levels of authority and competency, users will be more likely to trust a specialized robot than a fellow one, even if the specialization is arbitrarily assigned.

Our findings concerning the interaction effects demonstrated that users' perceived expertise was a *contingent* moderator (see Holbert & Park, 2020) in the relationship between perceived behavioral influence and users' modeling intentions. That is, for those who did not feel or only moderately felt that the robot was knowledgeable about garbage recycling, the more positive the perceived behavioral influence, the more willing they were to model the robot's behavior. However, for those who felt strongly about the robot's expertise in garbage recycling, the positive effect of perceived behavioral influence faded away. One possible explanation is that users became completely confident in and over-reliant on the robot's choices, which disincentivized them to shoulder responsibilities and learn the behavior themselves. This explanation can be supported by Hinds et al.'s (2004) finding that users felt less responsible when cooperating with a supervisor robot compared to a peer or subordinate robot.

5.3. *Theoretical and practical implications*

The findings of this study can spark pivotal developments in theory construction. First, prior research on SCT has been dominantly focused on users' vicarious learning in the contexts of mass and interpersonal communication. By discussing the ramifications of behavioral outcomes presented by social robots, this study is one of the first to put forward that SCT holds its predictive power in HRI. Furthermore, this study is pioneering in its contribution to the explanatory mechanism of SCT in HRI. Specifically, identification plays a necessary mediating role in social robots' modeling effects. Meanwhile, social presence is an imperative precursor to identification when social robots demonstrate modeled behavior without explicitly indicating the behavioral consequences.

Second, prior literature on TME has examined how social psychological findings, such as similarity attraction and reciprocity rules, can be replicated in HCI (Reeves & Nass, 1996). However, little research has been conducted regarding whether users could learn vicariously from social robots as they do from humans. This study endorses the heuristic value of TME and expands its scope by demonstrating that users do transfer interpersonal learning scripts to HRI and treat social robots as if they were human role models. Moreover, adding users' interpretation of robots' expertise to the equation, this study delineates the specific contingencies under which users' mimicking behavior is leveraged by their perceived behavioral outcomes.

The study also has rich practical implications. First, social robots can be used as role models to promote pro-social behavior like environment protection. Developers can consider designing positive behavioral consequences into robots' movements. If behavioral consequences cannot be delivered, designers can try embedding richer constellations of social cues and greater levels of communicative capabilities into robots to elevate users' social presence experience, as social presence is a necessary psychological component that indirectly leads to users' modeling tendencies when behavioral consequences are not observable. Even when social robots are used only for display in public spaces, demonstrating negative outcomes of a prototype behavior may increase users' learning interest and raise public awareness of the actions that have negative social impact.

Second, this study informs user experience design by demonstrating that even the arbitrary assignment of an instructor role to social robots can evoke users' modeling tendencies. Developers should therefore focus on using robots' messages or behavior to increase its perceived expertise and authority to promote pro-social behavioral change. Given the contingency of the interaction effects, developers should also be cautious in establishing robots' expertise in that users' over-reliance on robots' specialization may inversely decrease their own learning motivations.

Third, researchers should pay full attention to the potential ethical risks of manipulating social robots' behavior and social roles. For instance, companies or research institutions can develop ethical codes to prevent designers from utilizing robots to demonstrate anti-social messages or behavior. They should also carefully explain the potential influence of robots' behavior and messages when these robots are used in public sphere to guide users' behavior.

6. Conclusions and limitations

Turing's (1950) question "Can machines think?" marks a significant starting point for the academic discussion of artificial intelligence. While the past HCI paradigm has focused on how machines can be programmed, simulated, and trained to simulate human intelligence and autonomy, there has been a gap in the HRI literature regarding how humans may imitate machines and use machines to guide their own behavior. Therefore, this study inquires into this field and explores how social robots can be set as models to encourage users' modeling behavior. Although more research is needed to pave the way for a deeper understanding of social robots' modeling effects and their concomitant perils, the results of this study revealed optimistic prospects for utilizing robots' behavioral outcomes and social positions to promote pro-social behavior.

One of the limitations of the study is that behavioral outcomes were presented via robot *NAO*'s verbal messages. No immediate tangible outcome (e.g., environmental threats, workers' poor working conditions) was presented to participants. Unlike other SCT studies where punishments or rewards could be observed immediately, a lack of immediate incentives or disincentives might have inhibited participants from fully recognizing the importance of the behavioral consequences and thus rendered the effects of behavioral outcomes indistinct. Future research could adopt other forms of social robots, e.g., Pepper, on which a video screen can be used to display short clips about the vivid and real-time behavioral outcomes.

Second, due to technical difficulties, this study adopted a symbolic garbage recycling task. Future research can further augment the external validity and allow users and robots to sort real garbage into different recycling bins.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request. Some data can be accessed at https://osf.io/bhsr6/?view_only=e148010832b74c3d9246ffa45a1e7432

References

Ajzen, I., 1991. The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50, 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Andsager, J.L., Bemker, V., Choi, H.L., Torwel, V., 2006. Perceived similarity of exemplar traits and behavior: Effects on message evaluation. Commun. Res. 33 (1), 3–18. https://doi.org/10.1177/0093650205283099.

Bandura, A., 1977. Social learning theory. Prentice-Hall, Englewood Cliffs, NJ.

Bandura, A., 2009. Social cognitive theory of mass communication. In: Bryant, J., Oliver, M.B. (Eds.), Media Effects: Advances in Theory and Research, 2nd ed. Lawrence Erlbaum, Mahwah, NJ, pp. 94–124.

Bandura, A., Ross, D., Ross, S., 1961. Transmission of aggression through imitation of aggressive models. J. Abnorm. Soc. Psychol. 63, 575–582. https://doi.org/10.1037/h0045925.

Barco, A., de Jong, C., Peter, J., Kühne, R., & van Straten, C. L. (2020). Robot Morphology and Children's Perception of Social Robots: An Exploratory Study. *Proceedings of the Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 125-127. https://doi.org/10.1145/3371382.3378348.

Bevan, C., & Fraser, D. S. (2015). Shaking hands and cooperation in tele-present human-robot negotiation. *Proceedings of 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 247-254. https://doi.org/10.1145/2696454.2696490.

Čaić, M., Avelino, J., Mahr, D., Odekerken-Schröder, G., Bernardino, A., 2020. Robotic versus human coaches for active aging: An automated social presence perspective. Int. J. Soc. Robot. 12, 867–882. https://doi.org/10.1007/s12369-018-0507-2.

Carolus, A., Binder, J.F., Muench, R., Schmidt, C., Schneider, F., Buglass, S.L., 2019. Smartphones as digital companions: Characterizing the relationship between users and their phones. New Media Soc. 21 (4), 914–938. https://doi.org/10.1177/1461444818817074.

Castellano, G., De Carolis, B., D'Errico, F., Macchiarulo, N., Rossano, V., 2021. PeppeRecycle: Improving Children's Attitude Toward Recycling by Playing with a Social Robot. Int. J. Soc. Robot. 13 (1), 97–111. https://doi.org/10.1007/s12369-021-00754-0.

Chartrand, T.L., Bargh, J.A., 1999. The chameleon effect: the perception–behavior link and social interaction. J. Pers. Soc. Psychol. 76 (6), 893. https://doi.org/10.1037/0022-3514.76.6.893.

Cho, E., Molina, M.D., Wang, J., 2019. The effects of modality, device, and task differences on perceived human likeness of voice-activated virtual assistants. Cyberpsychol. Behav. Soc. Netw. 22 (8), 515–520. https://doi.org/10.1089/cyber.2018.0571.

Cohen, J., 2001. Defining identification: A theoretical look at the identification of audiences with media characters. Mass Commun. Soc. 4 (3), 245–264. https://doi.org/10.1207/S15327825MCS0403_01.

Compeau, D.R., Higgins, C.A., 1995. Computer self-efficacy: Development of a measure and initial test. MIS Q. 19 (2), 189.

Dijksterhuis, A., Bargh, J., 2001. The perception-behavior expressway: Automatic effects of social perception on social behavior. Adv. Exp. Soc. Psychol. 33, 1–40. https://doi.org/10.1016/S0065-2601(01)80003-4.

Duffy, B.R., 2003. Anthropomorphism and the social robot. Rob. Auton. Syst. 42 (3–4), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3.

Edwards, A., Edwards, C., Spence, P.R., Harris, C., Gambino, A., 2016. Robots in the classroom: Differences in students' perceptions of credibility and learning between "teacher as robot" and "robot as teacher". Comput. Hum. Behav. 65, 627–634. https://doi.org/10.1016/j.chb.2016.06.005.

Edwards, C., Edwards, A., Stoll, B., Lin, X., Massey, N., 2019. Evaluations of an artificial intelligence instructor's voice: Social Identity Theory in human-robot interactions. Comput. Hum. Behav. 90, 357–362. https://doi.org/10.1016/j.chb.2018.08.027.

Edwards, A. (2018). Animals, humans, and machines: Interactive implications of ontological classification. In A. L. Guzman (Ed.), *Human-Machine Communication: Rethinking communication, technology, and ourselves* (pp. 29–50). https://doi.org/10.3726/b14399.

Eyssel, F., Kuchenbrandt, D., 2012. Social categorization of a social robots: Anthropomorphism as a function of group membership. Br. J. Soc. Psychol. 51, 724–731. https://doi.org/10.1111/j.2044-8309.2011.02082.x.

Fadelli, I. (2021, February 11). *HuggieBot 2.0: A soft and human-size robot that hugs users on request.* Tech Xplore. https://techxplore.com/news/2021-02-huggiebot-soft-human-size-robot-users.html.

Fox, J., Bailenson, J.N., 2009. Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. Media Psychol. 12 (1), 1–25. https://doi.org/10.1080/15213260802669474.

Fox, J., Gambino, A., 2021. Relationship Development with Humanoid Social Robots: Applying Interpersonal Theories to Human-Robot Interaction. Cyberpsychol. Behav. Soc. Netw. 24 (5), 294–299. https://doi.org/10.1089/cyber.2020.0181.

Garrett, R.K., Danziger, J.N., 2008. Disaffection or expected outcomes: Understanding personal Internet use during work. J. Comput.-Mediat. Commun. 13 (4), 937–958. https://doi.org/10.1111/j.1083-6101.2008.00425.x.

Gong, L., 2008. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Comput. Hum. Behav. 24 (4), 1494–1509. https://doi.org/10.1016/j.chb.2007.05.007.

Groom, V., Nass, C., 2007. Can robots be teammates? Benchmarks in human–robot teams. Interact. Stud. 8 (3), 483–500. https://doi.org/10.1075/is.8.3.10gro.

Guzman, A.L., 2020. Ontological boundaries between humans and computers and the implications for Human-Machine Communication. Human-Machine Communication 1, 37–54. https://doi.org/10.30658/hmc.1.3.

Hayes, A.F., 2013. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. Guilford Press.

Hayes, A.F., Preacher, K.J., 2014. Statistical mediation analysis with a multicategorical independent variable. Br. J. Math. Stat. Psychol. 67 (3), 451–470. https://doi.org/10.1111/bmsp.12028.

Hendricks, A., 2002. Examining the effects of hegemonic depictions of female bodies on television: A call for theory and programmatic research. Crit. Stud. Media Commun. 19 (1), 106–123. https://doi.org/10.1080/07393180216550.

Hinds, P.J., Roberts, T.L., Jones, H., 2004. Whose job is it anyway? A study of human-robot interaction in a collaborative task. Human-Computer Interaction 19 (1–2), 151–181.

Hoffmann, L., Derksen, M., & Kopp, S. (2020, March). What a Pity, Pepper! How Warmth in Robots' Language Impacts Reactions to Errors during a Collaborative Task. *Proceedings of the Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 245-247. https://doi.org/10.1145/3371382.3378242.

Hoffmann, L., Krämer, N.C., Lv, C., 2021. The persuasive power of robot touch. Behavioral and evaluative consequences of non-functional touch from a robot. PLoS One 16 (5). https://doi.org/10.1371/journal.pone.0249554.

Holbert, R.L., Park, E., 2020. Conceptualizing, organizing, and positing moderation in communication research. Commun. Theory 30 (3), 227–246. https://doi.org/10.1093/ct/qtz006.

Horstmann, A.C., Bock, N., Linhuber, E., Szczuka, J.M., Straßmann, C., Krämer, N.C., 2018. Do a robot's social skills and its objection discourage interactants from switching the robot off? PLoS One 13 (7), 1–25. https://doi.org/10.1371/journal.pone.0201581.

International Society for Presence Research (ISPR). (2000). The concept of presence: Explication statement. ISPR. https://smcsites.com/ispr/.

Jung, Y., Lee, K.M., 2004. Effects of physical embodiment on social presence of social robots. Proceedings of PRESENCE 80–87.

Karutz, C.O., Bailenson, J.N., 2015. Immersive virtual environments and the classrooms of tomorrow. In: Sundar, S.S. (Ed.), the Handbook of the Psychology of Communication Technology. Wiley, pp. 290–310.

Kim, J., Park, H.S., 2011. The effect of uniform virtual appearance on conformity intention: Social identity model of deindividuation effects and optimal distinctiveness theory. Comput. Hum. Behav. 27 (3), 1223–1230. https://doi.org/10.1016/j.chb.2011.01.002.

Koh, Y.J., Sundar, S.S., 2010. Effects of specialization in computers, web sites, and web agents on e-commerce trust. Int. J. Hum Comput Stud. 68 (12), 899–912. https://doi.org/10.1016/j.ijhcs.2010.08.002.

LaRose, R., Mastro, D., Eastin, M.S., 2001. Understanding Internet usage: A social-cognitive approach to uses and gratifications. Soc. Sci. Comput. Rev. 19 (4), 395–413. https://doi.org/10.1177/089443930101900401.

Lee, K.M., 2004. Presence, explicated. Commun. Theory 14 (1), 27–50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x.

Lee, N., Kim, J., Kim, E., Kwon, O., 2017. The influence of politeness behavior on user compliance with social robots in a healthcare service setting. Int. J. Soc. Robot. 9 (5), 727–743. https://doi.org/10.1007/s12369-017-0420-0.

Lee, E.J., Nass, C., 2002. Experimental tests of normative group influence and representation effects in computer-mediated communication: When interacting via computers differs from interacting with computers. Hum. Commun. Res. 28 (3), 349–381. https://doi.org/10.1111/j.1468-2958.2002.tb00812.x.

Lee, K.M., Park, N., Song, H., 2005. Can a robot be perceived as a developing creature? Effects of a robot's long-term cognitive developments on its social presence and people's social responses toward it. Hum. Commun. Res. 31 (4), 538–563. https://doi.org/10.1111/j.1468-2958.2005.tb00882.x.

Lee, K.M., Peng, W., Jin, S.A., Yan, C., 2006. Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. J. Commun. 56 (4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x.

Li, J., Ju, W., Reeves, B., 2017. Touching a mechanical body: Tactile contact with body parts of a humanoid robot is physiologically arousing. Journal of Human-Robot Interaction 6 (3), 118–130.

Li, D., Rau, P.P., Li, Y., 2010. A cross-cultural study: Effect of robot appearance and task. Int. J. Soc. Robot. 2 (2), 175–186. https://doi.org/10.1007/s12369-010-0056-9.

Liew, T.W., Tan, S.M., 2018. Exploring the effects of specialist versus generalist embodied virtual agents in a multi-product category online store. Telematics Inform. 35 (1), 122–135. https://doi.org/10.1016/j.tele.2017.10.005.

Lombard, M., Ditton, T., 1997. At the heart of it all: The concept of presence. *Journal of Computer-Mediated*. Communication 3 (2), JCMC321. https://doi.org/10.1111/j.1083-6101.1997.tb00072.x.

Lombard, M., Xu, K., 2021. Social responses to media technologies in the 21st century: The media are social actors paradigm. Human-Machine Communication 2, 29–55. https://doi.org/10.30658/hmc.2.2.

Looije, R., Neerincx, M.A., Cnossen, F., 2010. Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. Int. J. Hum Comput Stud. 68 (6), 386–397. https://doi.org/10.1016/j.ijhcs.2009.08.007.

Macrae, C.N., Johnston, L., 1998. Help, I need somebody: Automatic action and inaction. Soc. Cogn. 16 (4), 400–417. https://doi.org/10.1521/soco.1998.16.4.400.

Mutlu, B., & Forlizzi, J. (2008). Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. *Proceedings of 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction.* 287-294. https://doi.org/10.1145/1349822.1349860.

Nabi, R.L., Clark, S., 2008. Exploring the limits of social cognitive theory: Why negatively reinforced behaviors on TV may be modeled anyway. J. Commun. 58 (3), 407–427. https://doi.org/10.1111/j.1460-2466.2008.00392.x.

Nass, C., Moon, Y., 2000. Machines and mindlessness: Social responses to computers. J. Soc. Issues 56 (1), 81–103. https://doi.org/10.1111/0022-4537.00153.

Nass, C., Fogg, B.J., Moon, Y., 1996. Can computers be teammates? Int. J. Hum Comput Stud. 45 (6), 669–678. https://doi.org/10.1006/ijhc.1996.0073.

Nass, C.I., Lombard, M., Henriksen, L., Steuer, J., 1995. Anthropocentrism and computers. Behavior & Information Technology 14 (4), 229–238. https://doi.org/10.1080/01449299508914636.

Nomura, T., Kanda, T., Suzuki, T., 2006. Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. AI & Soc. 20 (2), 138–150. https://doi.org/10.1007/s00146-005-0012-7.

Reeves, B., Nass, C., 1996. The media equation: How people treat computers, television, and new media like real people. Cambridge University Press.

Reeves, B., Yeykelis, L., Cummings, J.J., 2016. The use of media in media psychology. Media Psychol. 19 (1), 49–71. https://doi.org/10.1080/15213269.2015.1030083.

Reynolds, M., 2017. Facebook AI bot learns to mimic human reactions by watching Skype conversations. New Sci. https://www.newscientist.com/article/2146294-facebook-ai-learns-human-reactions-after-watching-hours-of-skype/.

Sah, Y. J., Yoo, B., & Sundar, S. S. (2011). Are specialist robots better than generalist robots. *Proceedings of the 6th international conference on Human-robot interaction*, 241-242. https://doi.org/10.1145/1957656.1957751.

Salem, M., Eyssel, F., Rohlfing, K., Kopp, S., Joublin, F., 2013. To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability. Int. J. Soc. Robot. 5 (3), 313–323. https://doi.org/10.1007/s12369-013-0196-9.

Searle, J.R., 1980. Minds, brains, and programs. Behav. Brain Sci. 3 (3), 417–424. https://doi:10.1017/S0140525X00005756.

Shaw-Garlock, G., 2009. Looking forward to sociable robots. Int. J. Soc. Robot. 1 (3), 249–260. https://doi.org/10.1007/s12369-009-0021-7.

Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2006). Interactive humanoid robots for a science museum. *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. 305-312. https://doi.org/10.1145/1121241.1121293.

Straten, C.L.V., Peter, J., Kühne, R., Barco, A., 2020. Transparency about a robot's lack of human psychological capacities: effects on child-robot perception and relationship formation. ACM Transactions on Human-Robot Interaction (THRI) 9 (2), 1–22. https://doi.org/10.1145/3365668.

Suchman, L., 2007. Human-machine reconfigurations: Plans and situated actions. Cambridge University Press.

Takayama, L., Groom, V., Nass, C., 2009. I'm sorry, Dave: i'm afraid i won't do that: social aspects of human-agent conflict. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2099–2108. https://doi.org/10.1145/1518701.1519021.

Tao, C.C., Bucy, E., P., 2007. Conceptualizing media stimuli in experimental research: Psychological versus attribute-based definitions. Hum. Commun. Res. 33, 397–426. https://doi.org/10.1111/j.1468-2958.2007.00305.x.

Thunberg, S., & Ziemke, T. (2020). Are people ready for social robots in public spaces? In *ACM/IEEE International Conference on Human-Robot Interaction* (pp. 482-484).

Turing, A.M., 1950. Computing machinery and intelligence. Mind 59, 433–460. https://doi.org/10.1093/mind/LIX.236.433.

Webb, T., Sheeran, P., 2006. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychol. Bull. 132, 249–268. https://doi.org/10.1037/0033-2909.132.2.249.

Westerman, D., Cross, A.C., Lindmark, P.G., 2019. I believe in a thing called bot: Perceptions of the humanness of "chatbots". Commun. Stud. 70 (3), 295–312. https://doi.org/10.1080/10510974.2018.1557233.

Wheeless, L.R., Grotz, J., 1977. The measurement of trust and its relationship to self-disclosure. Hum. Commun. Res. 3 (3), 250–257. https://doi.org/10.1111/j.1468-2958.1977.tb00523.x.

Xu, K., 2019. First encounter with robot Alpha: How individual differences interact with vocal and kinetic cues in users' social responses. New Media Soc. 21 (11–12), 2522–2547. https://doi.org/10.1177/1461444819851479.

Xu, K., Liao, T., 2020. Explicating cues: A typology for understanding emerging media technologies. Journal of Computer-Mediated Communication 25, 32–43. https://doi.org/10.1093/jcmc/zmz023.

Xu, K., Lombard, M., 2017. Persuasive computing: Feeling peer pressure from multiple computer agents. Comput. Hum. Behav. 74, 152–162. https://doi.org/10.1016/j.chb.2017.04.043.

You, S., & Robert, L. P. (2018, March). Human–robot similarity and willingness to work with a robotic co-worker. In *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 251-260). IEEE. https://doi.org/10.1145/3171221/3171281.

D. Zanatto M. Patacchiola J. Goslin S. Thill A. Cangelosi Do humans imitate robots? An investigation of strategic social learning in human-robot interaction In 15th ACM/IEEE International Conference on Human-Robot Interaction (2020, March). IEEE 449 457 10.1145/3319502.3374776.

Zhao, S., 2006. Humanoid social robots as a medium of communication. New Media Soc. 8 (3), 401–419. https://doi.org/10.1177/1461444806061951.

Zimmerman, F.J., Ortiz, S.E., Christakis, D.A., Elkun, D., 2012. The value of social-cognitive theory to reducing preschool TV viewing: a pilot randomized trial. Prev. Med. 54 (3–4), 212–218. https://doi.org/10.1016/j.ypmed.2012.02.004.