6-2011

Effects of Appearance and Functions on Likability and Perceived Occupational Suitability of Robots

Sau-Lai LEE
Nanyang Technological University

Ivy Yee-Man LAU
Singapore Management University, ivylau@smu.edu.sg

Ying-Yi HONG
Nanyang Technological University

DOI: https://doi.org/10.1177/1555343411409829

Published in Journal of Cognitive Engineering and Decision Making, Volume 5, Issue 2, 2011, Pages 232-250
https://doi.org/10.1177/1555343411409829

Effects of Appearance and Functions on Likability and Perceived Occupational Suitability of Robots

Sau-lai Lee
Nanyang Technological University
Ivy Yee-man Lau
Singapore Management University
Ying-yi Hong
Nanyang Technological University

ABSTRACT: This article reports three experiments that examined the association between (a) appearances and perceived capabilities of robots, (b) appearance and capabilities of robots and liking for the robots, and (c) perceived capabilities of robots and judgments concerning their suitability for different occupations. In Experiment 1, the authors found that participants perceived human- and animal-like robots to have relatively more warmth-related (e.g., emotion) capabilities than machinelike robots have. In Experiment 2, the authors found that liking for robots was not affected by their human likeness or their having warmth or competence capabilities. In Experiment 3, participants generally thought that robots should have information-processing and communication capabilities more than sensory and emotion capabilities. More interestingly, participants considered robots with different capabilities to be suitable for different occupations, preferring robots with emotion capabilities more in occupations that require frequent interactions with humans than in occupations that do not.

KEYWORDS: human-robot interaction, mental models, attitude

Introduction

The robotic industry has been zealously testing robots as service providers in social settings, such as schools and hospitals (Carey & Markoff, 2010; Christensen & Pacchierotti, 2005). Although most people, especially younger people, hold a positive attitude toward robots and are open to the idea of having robots doing different tasks at home (Scopelliti, Giuliani, & Fornara, 2005; Young, Hawkins,

---

ADDRESS CORRESPONDENCE TO: Sau-lai Lee, Division of Psychology, Nanyang Technological University, 14 Nanyang Drive, Singapore 637332, slleeh@ntu.edu.sg.

Sharlin, & Igarashi, 2009), such a positive attitude is not unequivocal. For example, Japanese, who are well known for their wide acceptance of robots, were found to be as concerned as Chinese and Dutch participants about the possible bad influences of robots when they are used in a child care context (Kulic & Croft, 2005; Nomura, Suzuki, Kanda, & Kato, 2006). As such, the social contexts would affect people's attitudes toward robots; yet little research has addressed why this is the case and what factors underlie people's judgments. To fill this knowledge gap, the goal of the present research was to examine how the physical appearance of robots and their functions (social-emotional vs. task-oriented functions) affect their likability and perceived suitability for different occupations. This research has important implications for designing robots. It sheds light on questions such as the following: Would people expect certain capability from the physical appearance of a robot? Would people necessarily like human-looking (human-like) robots better than they would machine-looking (machine-like) robots? Would the human-like robots be preferred across different occupational contexts?

Goetz, Kiesler, and Powers (2003) have shown that robots of different appearance are perceived to be suitable for different occupations. Specifically, participants looked at images of robots of different appearance—some were more humanlike and some were more machinelike—and indicated their liking for the robots and the occupations they thought the robots were suitable for. The researchers found a significant Appearance × Job Nature interaction effect: Participants preferred machinelike robots to human-like robots for jobs that are more task oriented (e.g., lab assistant, customs inspector, soldier, security guard) but preferred humanlike to machinelike robots for jobs that are more people oriented (e.g., actress, drawing instructor, retail clerk, sales representative). Goetz and her associates proposed the matching hypothesis to explain their findings. They argue that robots of different appearance are preferred for different kind of tasks because people associate different characteristics to different physical appearance. Accordingly, humanlike robots convey animistic and anthropomorphic cues that activate people's perception of lifelikeness in them; the perception then leads people to attribute personality and emotional capabilities to the robots. Consequently, humanlike robots are deemed to be suitable for people-oriented jobs. On the other hand, machinelike robots, because they possess no animistic and anthropomorphic cues, are judged to be suitable for jobs that are task oriented and that involve little human contact. Unfortunately, Goetz et al. (2003) have only hypothesized but have not actually measured people's perception of robots' personality and emotional capabilities in their study and thus have not tested the proposed matching hypothesis empirically.

Subsequently, Powers and Kiesler (2006) proposed the mental model theory to further explain how the appearance of a robot may affect the perception of its suitability for different occupations. They suggested that when trying to understand robots, people apply similar stereotypes, social heuristics, and habitual social responses that they use to understand humans. These familiar tools for understanding humans are called forth by the presence of well-learned social cues. For example, a female voice or a pair of red lips on a robot calls for the social category *female* as well as the corresponding social knowledge associated with the category. According to the knowledge activation theory (Higgins, 1996), these activated constructs are then applied to understand the robot. Furthermore, when applicable to the situation, the activated constructs also guide perceivers' perceptions and behaviors in relation to the robot. Thus, depending on the constructs activated by the physical cues of the robot, people may attribute different knowledge, values, dispositions, and capabilities to the robot and react and behave accordingly.

To summarize, studies have shown that the appearance of robots affects people's perception of them. People ascribe different mental models to a robot according to the cues presented in the appearance of the robot, and the mental model evoked concerning the knowledge and capabilities of the robot affects people's judgment of its suitability for different occupations.

The Present Study

Although the aforementioned research has provided a framework to understand the relation between a robot's appearance, evoked mental model, and perceptions and judgments toward a robot, a number of questions remain unanswered. First, would people infer certain capabilities (functions) for robots with certain physical appearance? Research thus far has mainly focused on mental models of personality and social categories (e.g., the robot looks like a female), but no study has investigated the influence of mental models of capabilities of robots. Since capabilities are the ground for a lot of higher-level characteristics (e.g., a robot cannot be talkative if it cannot decode language and speak) and capabilities are highly correlated with a robot's suitability for different work settings (e.g., a robot cannot be a street cleaner if it cannot move), it is therefore important to understand how people infer the capabilities of a robot on the basis of its appearance.

Second, what is the relative contribution of a robot's appearance and its capabilities to its general likability to people? Since the inferred capabilities of the robots used in previous mental model research was not controlled, it was not possible to ascertain whether people's liking, or dislike, of a robot was attributable to its appearance or its inferred capabilities or lack thereof. The present study attempted to tease apart the independent contribution of these two factors.

Third, how does learning about people's mental model of a robot's capabilities help researchers understand people's judgment of the suitability of the robot for different occupations? One of the aims of the present research was to supplement previous works by providing a framework that could help answer this question.

Research has found two dimensions in intergroup perceptions (Fiske, Cuddy, Glick, & Xu, 2002), namely, warmth and competence. *Warmth* refers to perceptions of kindness, honesty, helpfulness, and trustworthiness of an entity. *Competence* refers to perceptions of intelligence, capability, skillfulness, and competitiveness of an entity. In a broader sense, warmth depicts a general tendency to be people oriented, whereas competence depicts a general tendency to be more task oriented. It has been shown that these two dimensions can be used to understand the perception of both human groups and social organizations (Aaker, Vohs, & Mogilner, 2010). These two dimensions are orthogonal to each other and can be crossed to form four quadrants. Exemplars of social groups in each quadrant are rich people, who are perceived as high on competence but low on warmth; housewives, who are perceived as high on warmth but low on competence; poor people, who are perceived as low on both competence and warmth; and professionals, who are perceived as high on both competence and warmth (Fiske et al., 2002).

In this article, we propose that the warmth and competence dimensions could be incorporated into the matching hypothesis and mental model theory to understand people's perception of robots and the mental models that robots' appearance can activate, thereby predicting the perceived suitability of robots for different occupations. Specifically, in line with the matching hypothesis, humanlike robots possess human features (such as eyes or mouths) that can afford warmth-related capabilities (e.g., recognizing and expressing emotions), whereas machinelike robots possess features (such as mechanical parts) that can afford task competence-related capabilities (e.g., computation and memory). Thus, humanlike robots should call up mental models related to the warmth dimension and be perceived as higher on warmth than on competence, whereas machinelike robots should call up mental models related to the competence dimension and be perceived as higher on competence than on warmth. Moreover, different occupations require different skills—some occupations are more people focused, and some are more task focused—therefore it is reasonable to expect that robots varying on perceived warmth and competence also should be perceived to be differentially suitable for occupations requiring different levels of warmth versus competence. The most optimal situation is when the perception of robot capabilities matches with the perceived need of the occupation. Thus, people should be more approving of a robot that is perceived to be high on warmth (or competence) to perform a people-focused (or task-focused) job.

To test these ideas, we have conducted three experimental studies to test six hypotheses (as summarized in Table 1). Experiment 1 aimed to test the systematic relation between a robot's appearance and the mental models activated, especially those relating to the robot's capabilities. Participants were shown pictures of a variety of machinelike, animal-like, or humanlike robots and were asked to guess the capabilities of each robot. Because human features are more salient in humanlike robots and electronic and mechanical features are more prominent in machinelike robots, we predicted that participants would infer more warmth-related capabilities (e.g., can decode emotion and have sensation) for the humanlike robots than for the machinelike robots (Hypothesis 1), and they would infer more competence-related capabilities (e.g., information-processing capabilities) for the machinelike robots than for the humanlike robots (Hypothesis 2).

Experiment 2 aimed to test the independent contribution of a robot's appearance and functions to its likability. In past research, people were asked to rate their perception of robots based on the robots' appearance. As people may infer certain capabilities from a robot's appearance (as shown in Experiment 1), it is

TABLE 1. Hypotheses Tested in Each Experiment

Experiment	Hypothesis	Hypothesis Supported?	Findings
Experiment 1	Hypothesis 1: More warmth-related capabilities will be inferred for the humanlike than for the machinelike robots.	Yes	Participants perceived humanlike and animal-like robots to have more warmth-related (i.e., emotion and communication) capabilities than machinelike robots
Experiment 1	Hypothesis 2: More competence-related capabilities will be inferred for the machinelike robots than for the humanlike robots.	No	Participants perceived the three types of robots to have a similar number of competence-related (i.e., information-processing and cognition) capabilities.
Experiment 2	Hypothesis 3: Humanlike robots are more likable than robots that are not humanlike.	No	Humanlike robots were not more likable than robots that were not humanlike.
Experiment 2	Hypothesis 4: Robots that are perceived to possess more warmth-related capabilities should be liked more than those perceived to possess competence-related capacities.	No (a trend consistent with the hypothesis)	Robots that possessed warmth-related capabilities were not rated as more likable than robots that possessed competence-related capabilities only.
Experiment 3	Hypothesis 5: Robots that are designed to have more warmth-related functions would be seen as more suitable for people-oriented jobs.	Yes	Emotion-related capabilities were a significant predictor of robots' suitability to people-oriented occupations.
Experiment 3	Hypothesis 6: Robots that are designed to have more competence-related functions would be seen as more suitable for task-oriented jobs.	Yes	Information-processing capabilities were a significant predictor of the robots' suitability to task-oriented occupations.

difficult to discern whether people's liking of the robot was attributable to its appearance or its inferred capabilities (as part of the activated mental model). By manipulating the appearance and functions of a robot independently in Experiment 2, we were able to test the relative effects of appearance and functions in predicting the likability of robots. Specifically, according to the literature on anthropomorphism (see Epley, Waytz, & Cacioppo, 2007, for a review), people should like a humanlike robot more than they do a robot that is not humanlike because of affordability of the humanlike robot to form social connections with humans (Hypothesis 3); for the same reason, people should also like robots that possess social-emotional capabilities more than they do robots that possess task-oriented capabilities (Hypothesis 4).

Experiment 3 aimed to test whether participants would ascribe different capabilities to a robot according to its appearance and, on the basis of the ascribed capabilities, how they would judge its suitability for different occupations. In this study, we asked participants to select the capabilities of four robots and to rate their suitability for different kinds of occupations. We expected that warmth-related capabilities of a robot would predict its suitability for more people-oriented jobs (Hypothesis 5), whereas competence-related capabilities would predict its suitability for more task-oriented jobs (Hypothesis 6).

Experiment 1

Method

Participants. Thirty-three undergraduates (14 male, 19 female) from a local university in Singapore, all proficient in English, participated in exchange for course credit. The age of the participants ranged from 18 to 24 ($M = 20.55$, $SD = 1.54$). They participated in the experiment in groups of no more than 6.

Design. This experiment involved a 3 (type of robot: machine-, animal- and human-like) × 7 (type of function: communication, cognition, emotions, information processing, movement, motivation, and sensory) design, with type of robot as between-subjects variable and function as within-subjects variable. Participants were asked to guess whether each robot was capable of performing the seven groups of functions.

Procedure. Participants were randomly assigned to one of the three robot type conditions and were shown 10 pictures of either machinelike, animal-like, or humanlike robots on a PowerPoint slide show one at a time. The robots were selected by the authors on the basis of the following criteria: For humanlike robots, they must have at least two human features, for example, facial features, such as eyes, and body parts, such as limbs. For animal-like robots, they must resemble one kind of animal, for example, a cat or a dog. For machinelike robots, they must not have any aforementioned feature (usually, they are box-shaped and have wheels). Participants were then asked to check the possible capabilities of each robot from a list of 29 functions (see Table 2). There were seven categories of functions, namely, communication (e.g., speech production), senses (e.g., hearing), emotions (e.g., display emotions), information processing (e.g., information storage), movement (e.g., avoid obstacles), cognition (e.g., learning), and motivation (e.g., curiosity). Both communication and emotions were classified as warmth-related capabilities, as they were directly associated with the robots' capabilities to relate to people. The rest were classified as competence-related capabilities, as they were associated with work efficiency of the robots. The participants also rated their liking for each robot on a 7-point Likert-type scale ranging from

TABLE 2. Capabilities Used in Experiments 1and 3

Function Category (Number of Items in Each Category)	Capabilities
Communication (5)	Speech production; speech comprehension; ability to carry on conversation, give out signals, relate to people or other robots
Cognition (2)	Learning, problem solving
Emotions (5)	Ability to experience emotions, display emotions, decode human emotions, feel compassion, care for others
Information processing (2)	Computation, information storage
Motivation (5)	Ability to create, be curious, have ambition, lie, decode human intention
Movement (4)	Movement; movement with sensory capability and ability to respond accordingly (e.g., avoid obstacles); ability to transport objects, move like a human
Sense (6)	Sense of touch, smell, hearing, sight, temperature, taste

−3 (*dislike it very much*) to +3 (*like it very much*). This scale was designed to explore how people think of different kinds of robots in general. All the materials and instructions used in this and other experiments in the study were in English.

Results

When testing the hypotheses for this article, we used a significance level of .05, two tailed, as the criterion for rejecting the null hypothesis.

We calculated a ratio score for each of the seven functions by averaging the numbers of checks for the functions across the 10 robots. Take the communication function as an example. There were five possible capabilities. The ratio score for communication function was the average number of checks among the five possible capabilities across the 10 robots, divided by 5. As a result, each participant had seven ratio scores, one for each category of function; the higher the score, the greater capabilities the robots were perceived to have in that function.

To test our hypothesis that people ascribe different warmth versus competence functions to robots according to their appearance, we submitted the seven average scores to seven separate one-way ANOVAs, with type of robot as between-subjects factor. We did not perform a Function × Robot ANOVA because our sample size was relatively small (Rosech, 2010; Tabachnick, & Fidell, 2007). As shown in Figure 1, in general, participants assigned more information-processing and movement capabilities than motivation and emotion capabilities to the robots regardless of their appearance. In addition, participants thought that the humanlike and animal-like robots generally would have more functions than machinelike robots. Results from the ANOVA tests and post hoc analysis (using Tukey B) showed that out of the seven categories of functions, participants ascribed significantly more communication, $F(2, 30) = 4.89$, $p < .05$, partial $\eta^2 = .25$, and emotion, $F(2, 30) = 4.83$, $p < .05$, partial $\eta^2 = .24$, capabilities to both humanlike (communication, $M = .37$, $SD = .12$; emotion, $M = .13$, $SD = .10$) and

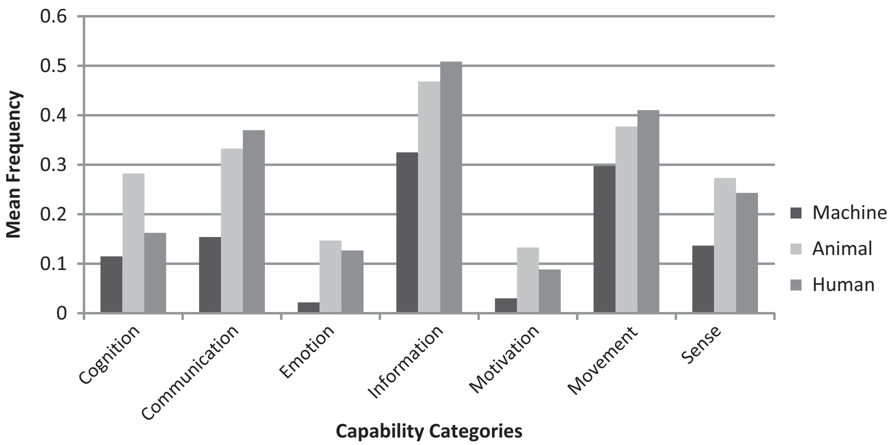

Figure 1. Mean endorsement frequency of each kind of capability.

animal-like (communication, *M* = .33, *SD* = .25; emotion, *M* = .15, *SD* = .14) robots than to machinelike (communication, *M* = .15, *SD* = .08; emotion, *M* = .02, *SD* = .02) robots. However, they ascribed a similar extent of capabilities to the remaining five categories of capabilities to the three kinds of robots.

Results supported Hypothesis 1: Participants ascribed more emotion and communication capabilities (warmth-related capabilities) to humanlike robots than to machinelike robots. They also ascribed more warmth-related capabilities to animal-like than to machinelike robots. This pattern of results would be predicted by the mental model theory, as the appearance of the humanlike and animal-like robots cued a more animistic mental model of the robots; therefore, accordingly, participants assumed human- and animal-like robots to be more likely to possess warmth-related capabilities. However, Hypothesis 2 was not supported, as participants did not ascribe more competence-related capabilities to machinelike than to humanlike robots. Participants ascribed similar movement and information-processing capabilities to all three types of robots. A possible explanation is that participants might have assumed that most robots would have some basic information-processing and movement capabilities. This possibility is reflected in the fact that information-processing and movement functions are the most frequently chosen capabilities. Finally, inconsistent with previous findings, in the present study, humanlike robots were not rated as more likable than the other two kinds of robots, $F(2, 30) = .41$, *ns*. Indeed, the three kinds of robots did not differ on their likability to the participants.

Experiment 2

Experiment 2 aimed to test the relative contribution of appearance and functions to the likability of a robot. We presented participants with robots of different appearance and at the same time manipulated the capabilities of the robots by telling one group of participants that the robots could decode and express emotions (warmth condition) and the other group that the robots could perform computation and solve problems (competence condition).

Method

Participants. Forty undergraduates from a local university in Singapore (20 male, 20 female) participated for course credit. The group consisted of 92.5% ethnic Chinese, 5.1% ethnic Indian, 2.6% ethnic Malay participants. They participated in the experiment in groups of no more than 6.

Design. This experiment involved a 4 (robot: Papero, Pearl, Asimo, Repliee Q2) × 2 (function: warmth vs. competence) design, with robot as within-subjects variable and function as between-subjects variable. The dependent variables were the likability and the human likeness of the robots.

Procedure. We selected pictures of four robots used in Experiment 1 as stimuli instead of using all 30 to avoid possible participant fatigue. The robots were selected to represent a range of human likeness and for their affordance of the functions manipulated in the study. From the least humanlike to the most humanlike, the chosen robots were Papero, which was approximately 15 inches tall, with separate torso and head and with two eyes (developed by NEC); Pearl, which was approximately 4 feet tall, with a clear head and torso and with eyes and a mouth, although it was run on wheels (developed by a group of researchers at Carnegie Mellon University); Asimo, a standing robot approximately 5 feet tall, with four limbs and separated head and torso (developed by Honda); and Repliee Q2, a seated robot that looked almost indistinguishable from a human, with full features and skin that looked like a real human's (developed by the Intelligent Robotic Laboratory at Osaka University). Participants were shown pictures of the robots along with a description of their capabilities: They were told that the robots possess either social emotional functions (the warmth condition) or computation and problem-solving functions (the competence condition). For example, half of the participants (randomly assigned) were told that Papero can express and decode emotions, and the other half were told that Papero can carry out complex computation and solve problems. The order of the presentation of the pictures was counterbalanced. Participants first rated the likability of the robots on a 7-point Likert-type scale from –3 (*dislike it very much*) to +3 (*like it very much*), and then rated the human likeness of the robots on a 7-point scale from 1 (*not humanlike at all*) to 7 (*very humanlike*).

Results

To ascertain the range of perceived human likeness of the robots, we compared the human likeness ratings across the robots using a one-way within-subjects ANOVA. The result showed a significant main effect of robot, $F(3, 111) = 98.43$, $p < .05$, partial $\eta^2 = .73$. A post hoc pairwise comparison using LSD showed that there was a linear increment in the human likeness from Papero (*M* = 2.05, *SD* = 1.23) to Pearl (*M* = 2.45, *SD* = 1.55) to Asimo (*M* = 2.87, *SD* = 1.84) to Repliee Q2 (*M* = 5.76, *SD* = 1.38), which was consistent with our ranking of the human likeness of these robots.

We tested the effect of appearance and function on likability of the robots by submitting the likability scores to a 4 (robot) × 2 (function) ANOVA, with robot as a within-subjects factor and function as a between-subjects factor. The interaction effect of Robot × Function was not significant, $F(3, 108) = .18$, *ns*. However, there was a significant main effect of robot, $F(3, 108) = 4.28$, $p < .05$, partial $\eta^2 = .11$. A post hoc pairwise comparison using LSD showed that the likability for Pearl (*M* = 3.32, *SD* = 1.67) was the lowest, followed by that for Repliee Q2 (*M* = 4.06, *SD* = 1.64), then for Asimo (*M* = 4.41, *SD* = 1.73), and finally, for Papero (*M* = 4.54, *SD* = 1.83) (see Figure 2). Liking for Repliee Q2, Asimo, and Papero were not significantly different. This result is inconsistent with Hypothesis 3, which predicted that the more humanlike a robot is, the more it would be liked.

The main effect of function was not significant, $F(1, 36) = 3.47$, $p = .07$; therefore, Hypothesis 4, which predicted that people would like robots that have warmth-related capabilities more than competence-related capabilities, was not supported. However, there was a trend that the same robots were rated as more likable when they could decode and recognize emotions (*M* = 4.38, *SD* = 0.94) than when they could process information (*M* = 3.79, *SD* = 0.94).

To understand the relation between the human likeness of a robot and participants' liking for it when it possessed warmth versus competence capabilities, we conducted generalized linear modeling (GLM), one for each robot, with its likability as dependent variable and its mean-centered human likeness and function as predictors. For all four robots, the main effects of human likeness and function were not significant; neither were the Human Likeness × Function interactions.

As a whole, the findings were inconclusive in terms of how appearance and functions affect people's liking of robots. Contrary to our predictions, humanlike robots were not liked more than machinelike robots; robots with warmth capabilities were not liked more than those with competence capabilities (although there was a slight advantage for robots with warmth capabilities). Furthermore, there was no interaction between the perceived human likeness of the robots and functions on their likability. Hypotheses 3 and 4 were not supported.

Experiment 3

Experiment 2 tested the effect of appearance and capabilities on the likability of robots. However, it is still unknown how these two variables may affect the perceived suitability of the robots for different occupations. To fill this knowledge gap, we asked participants first to rate the functions of a robot with a certain appearance and then to rate the suitability of the robot for various occupations. As we found in Experiment 1 that people would associate different functions with robots of different appearance (humanlike, animal-like, and machinelike), we can obtain a wide range of functions by presenting participants with robots of different appearance.

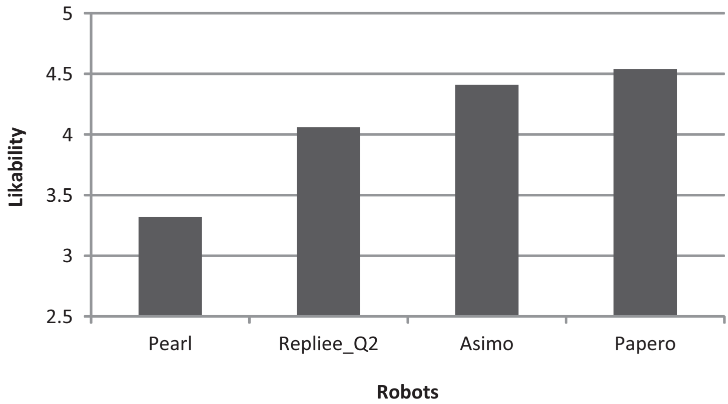

Figure 2. Mean likability of robots used in Experiment 2.

Subsequently, we can test how the appearance and the functions interact to predict participants' perceived suitability of the robots for different occupations.

As a secondary aim, the third experiment also examined whether participants' judgments would differ if the robot were to work alone or collaborate with humans on a job. Specifically, in Experiments 1 and 2, people were asked to think about the robots as independent entities; people's perception of the robots might be different if they were to consider them as collaborators at work. People might want a robot to be able to decode their internal states more when it is a coworker than when it is not related to them. Thus, to address these two issues, in Experiment 3, we manipulated the context within which participants judged the robots. Half of the participants were told to consider the presented robot an individual entity, and the other half of the participants were told to consider the robot a potential collaborator. We expected that relative to participants who considered the robot as an independent entity, participants who considered the robot as a potential collaborator would prefer more that it had emotional capabilities.

Method

Participants. Two hundred undergraduates from a local university in Singapore participated in this study for course credit. Their mean age was 20.63 (SD = 1.63). They participated in the experiment in groups of no more than 6.

Design. This experiment involved a 4 (robot: Aibo, Asimo, Papero, Repliee Q2) × 2 (context: individual vs. collaboration) design with both robot and context as between-subjects variables. The two dependent variables were the agreement between each robot and each kind of capability and the suitability of the robots for different occupations.

Procedures. The four robots used in this experiment were the same as in Experiment 2, except we replaced Pearl with Aibo (a doglike robot developed by Sony) to generalize the findings to animal-like robots. By way of introduction to the study, participants read that with the advancement of technology, robots are becoming more and more sophisticated and are including more and more capabilities. They were then asked to imagine that they were part of a project team that was helping to design the functions of a robot (a picture of the corresponding robot was shown to them) and that there was no constraint on the functions that it could perform. Participants were randomly assigned to view one of the four robots in one of the two instruction conditions. For the individual condition, participants were asked to consider the robot an individual entity. For the collaborator condition, participants were asked to consider the robot a future collaborator for humans in both work and nonwork settings. Then all participants were asked to indicate the extent to which they agreed that the robot should have each of the 29 specific capabilities from the seven capability categories used in Experiment 1. They rated their agreement on a 7-point Likert-type scale from –3 (*strongly disagree*) to +3 (*strongly agree*). Participants then rated the suitability of the robot for different occupational categories on a 7-point Likert-type scale from –3 (*very unsuitable*) to +3 (*very suitable*). Participants were presented with 23 occupational categories that were adopted from the Standard Occupational Classification of the U.S. Bureau of Labor Statistics (http://www.bls.gov/soc/soc_majo.htm). An example of each occupational category was provided to the participants.

Results

To reduce the number of function categories and occupational categories, we submitted both the capabilities agreement ratings and the occupation suitability ratings to two separate factor analyses. We selected the solution on the basis of the results of the scree plots. For capabilities, a three-factor solution was selected. The model explained 45.82% of the variance. An examination of capabilities with an eigenvalue of at least 2.41 (all except one variable has significant loading on only one factor) revealed that Factor 1 was related to emotion, motivation, and cognition capabilities; Factor 2 was related to different kinds of sensory capabilities; and Factor 3 was related to information-processing and communication capabilities. For occupation suitability, a three-factor solution that explained 44.40% of variance was selected. An examination of ratings with an eigenvalue of at least 1.90 showed that Factor 1 comprised mainly blue-collar occupations (e.g., butcher, carpenters, repairers), Factor 2 comprised mainly professional occupations (e.g., financial specialist, programmer, architects), and Factor 3 comprised mainly social and service-related occupations (e.g., counselor, hairdressers, teacher, designers). For each participant, we averaged the ratings for each of the three function and three occupation groups.

To examine whether the participants ascribed different capabilities to different kinds of robots in different judgment contexts, we submitted the three averaged function scores to a 4 (robot) × 2 (context) × 3 (function) ANOVA, with the first two variables as between-subjects factors and the last variable as within-subjects factor. The three-way interaction was not significant, $F(6, 384) = 0.67$, *ns*. However, there was a significant Robot × Function interaction, $F(6, 384) = 3.02$,

$p < .05$, partial $\eta^2 = .05$, and a significant Context × Function interaction, $F(2, 384) = 3.33$, $p < .05$, partial $\eta^2 = .02$. Our follow-up repeated-measure ANOVAs on the three function groups, $F(1, 49) \geq 47.82$ (significant for each robot), showed that for all robots except Aibo, participants in general judged that the robots should have information-processing and communication capabilities more than sensory capabilities and sensory capabilities more than emotion-, motivation-, and cognition-related capabilities (see upper panel of Figure 3). This pattern was similar to what we found in Study 1. A pairwise comparison showed that for Aibo, participants preferred it to have sensory and information-processing capabilities to the same extent.

As shown in the lower panel of Figure 3, with Aibo excluded, regardless of the judgment context, the relative ranking of the participants' ratings for the functions were about the same, with information processing and communication rated the highest, followed by sensory capability; the lowest ratings were given for emotion, motivation, and cognition. However, the participants on average judged that a robot should have significantly more emotion-, motivation-, and cognition-related functions when it was a potential collaborator than when it was an individual entity, $t(198) = -2.66$, $p < .05$.

Since the ratings for functions follow the same general trend across the three humanoid robots and the two judgment contexts, we combined the data from the two context conditions and the three robots in the subsequent analysis when we tried to predict the effect of functions on perceived occupation suitability. We regressed the suitability ratings for each of the three groups of occupations on the three groups of functions using GLM. First of all, all three models were significant, with values of $F(2, 146)$ of at least 10.50. As predicted in Hypothesis 6 (competence-related capabilities would predict a robot's suitability for more task-oriented jobs), the more the participants ascribed information-processing and communication capabilities (which are subsumed under competence-related capabilities) to the robots, the more they rated them as suitable for task-oriented blue-collar occupations, $\beta = .48$, $t(146) = 4.66$, $p < .05$, partial $\eta^2 = .13$. Sensory, $\beta = .06$, $t(146) = 0.72$, *ns*, and emotion-, motivation-, and cognition-related, $\beta = .10$, $t(146) = 1.30$, *ns*, capabilities could not predict the robots' suitability for blue-collar occupations.

For professional occupations, the more participants ascribed the warmth-related emotion , motivation, and cognition capabilities, $\beta = .24$, $t(146) = 3.13$, $p < .05$, partial $\eta^2 = .06$, and competence-related information-processing and communication capabilities, $\beta = .40$, $t(146) = 3.66$, $p < .05$, partial $\eta^2 = .08$, to the robots, the more they thought the robots would be suitable. Sensory capabilities could not predict suitability of the robots for professional occupations, $\beta = -.16$, $t(146) = -.16$, *ns*. The results are consistent with previous research showing that professionals are perceived as high on both warmth and competence (Fiske et al., 2002).

Finally, for social and service-related occupations, we found that only emotion, motivation, and cognition capabilities could predict the robots' suitability, $\beta = .40$, $t(146) = 5.90$, $p < .05$, partial $\eta^2 = .19$, whereas sensory, $\beta = -.06$, $t(146) = -0.79$, *ns*, and information-processing and communication capabilities could not, $\beta = -.03$, $t(146) = -0.34$, *ns*. The results are consistent with Hypothesis 5, as

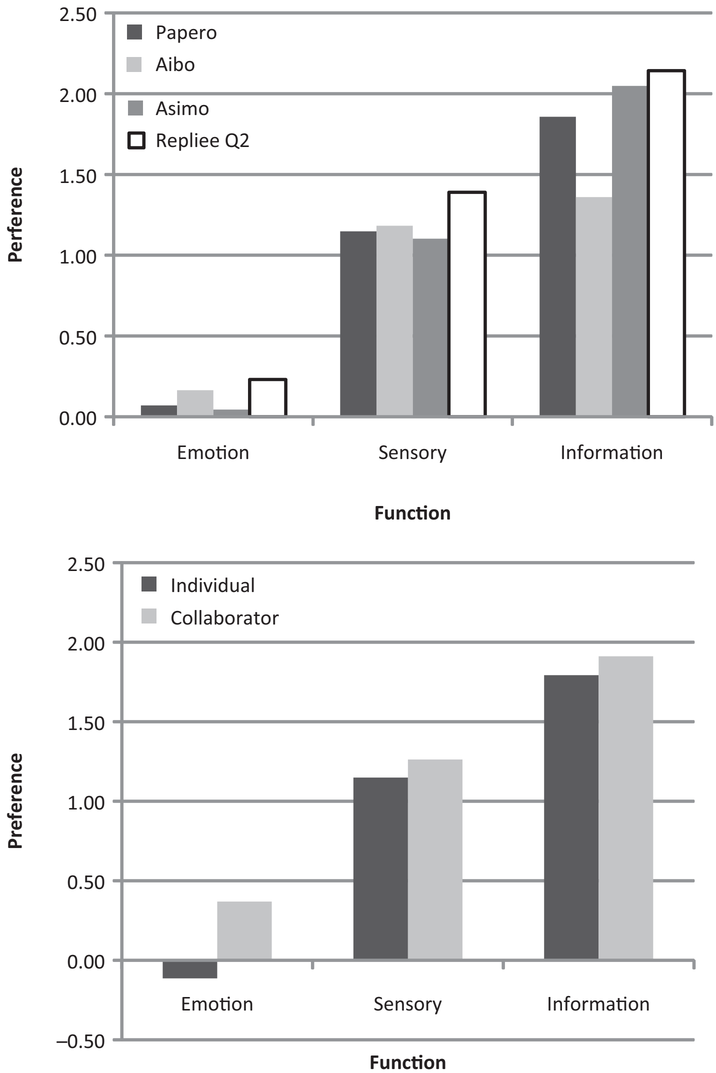

Figure 3. Mean preference ratings of functions for different kinds of robots (upper panel) and mean preference ratings of functions in different judgment contexts (lower panel).

social and service-related occupations are more people focused; therefore, the more the robots have emotion, motivation, and cognition capabilities, the more suitable they were perceived to be for these occupations.

General Discussion

The mental model theory proposes that different features (e.g., appearance, size, and voice) of robots cue different mental models about the robots. The present research aimed to test (a) whether the mental models for the capabilities of robots of different appearance are systematically different, (b) how appearance and capabilities of robots conjointly affect people's liking of robots, and (c) how robots' appearance and perceived capabilities affect people's perception of their suitability for different occupations.

To answer the first question, Experiment 1 aimed to test whether the perceived capabilities for machinelike, animal-like, and humanlike robots were different. In terms of quantitative difference, participants perceived human- and animal-like robots to have more capabilities than machinelike robots. In terms of quality, consistent with what we have predicted, participants perceived both the human- and animal-like robots to have more warmth-related capabilities, such as communication and emotion capabilities, than the machinelike robots. Contrary to our prediction, however, people did not perceive the machinelike robots to have more competence-related capabilities, such as information processing, than human- or animal-like robots. These findings by and large are consistent with the mental model theory: People apply social heuristics that they have learned in their daily life to estimate the capabilities of the robots. As in everyday-life situations, humans and animals do possess more emotion and communication capabilities than do machines, and thus they would be more capable to show warmth than would machines. However, regarding competence-related capabilities, people seem to assume that robots, regardless of their appearance, possess a high level of such capabilities by default.

To answer the second question, we conducted Experiment 2, in which we manipulated the functions. We told participants that robots of different appearance can either express emotions (warmth condition) or can perform computation and solve problems (competence condition). Unfortunately, the findings were inconclusive: Participants' liking of the robots was not systematically related to the robots' appearance (human likeness). That is, contrary to our prediction, participants did not necessarily like robots with high human likeness more than they did those with low human likeness. Also, there was only a trend that participants liked the robots more in the warmth than in the competence condition. These results need to be explored further with a larger sample size.

To answer the third question, Experiment 3 tested whether people's judgment of a robot's suitability for an occupation was affected by its appearance and capabilities that people ascribed to it. Consistent with our prediction, the more the participants ascribed competence-related capabilities (information-processing and communication capabilities) to a robot, the more they rated the robot as suitable for blue-collar occupation. Also, the more the participants ascribed warmth-related capabilities (emotion, motivation, and cognition) to a robot, the more they rated the robot as suitable for social and service-related occupations. Both warmth- and competence-related capabilities were positively correlated with the robots' suitability for professional types of occupations. This finding is consistent with the matching hypothesis (Goetz et al., 2003; Torrey, Powers, Marge, Fussell, & Kiesler, 2006), which argues that robots of different appearance are preferred for different kinds of tasks. We have extended the hypothesis to the realm of robot capabilities by showing that robot capabilities also affect the judgment of suitability for different kinds of tasks.

Implications for Designing Robots

The present research provides an important framework to inform design of robots for different work settings. Although most of the existing findings have shown that in general, humanlike robots are better liked than machinelike robots, we argued that this finding may not necessarily be the case for robots in all work settings. Findings from the present research suggest that the design of the appearance and functions of a robot should match the nature of the occupation that it is assigned to. Specifically, Experiment 1 revealed that people have mental models regarding the functions of robots of different appearance. Thus, people may form expectations of the functions of a robot on the basis of its appearance. When the actual capabilities of a robot are less than what a user expects from its appearance (e.g., a humanlike robot that cannot talk), the user may experience disappointment. When the actual capabilities of a robot are better than what a user expects from its appearance (e.g., a machinelike robot that can talk), the user may experience surprise. Thus, when roboticists design the appearance of a robot, it would be desirable to design features that are commensurate with the capability expectations that arise from the appearance.

Furthermore, a match between functions of the robots and occupational demands is also important. For example, warmth capabilities should be installed in robots when they are needed in work contexts involving interactions with humans. In general, for people-oriented jobs, a robot designed with humanlike appearance and warmth capabilities may enhance user acceptance and satisfaction. For task-oriented jobs, humanlike appearance and warmth capabilities may be less crucial.

That being said, a designer has to be cautious about humanlike robots, especially those that bear a high resemblance to human (humanoids). Specifically, humanlike robots did not necessarily evoke greater likability. In Experiments 1 and 2, we did not find a positive relation between the human likeness of the robots and their likability. In fact, in Experiment 2, the most well-liked robot was the least humanlike (Papero). This result is inconsistent with the prediction from past literature on anthropomorphism (Bartneck, Kanda, Mubin, & Al Mahmud, 2009). However, it is unlikely that our results were attributable to an inadequate sample of robots used in the experiments, as there were 10 robots for each of the three categories in Experiment 1. Instead, the phenomenon of "the uncanny valley" in robotics may shed some light on this finding. The uncanny valley refers to the situation in which human observers feel revulsion when robots or other artifacts look and act almost like real human beings (MacDorman, 2006). This phenomenon suggests that the relationship between human likeness and likability may not be linear; an extremely high level of human likeness (as in humanoids) could result in a lower level of likability. Designers may need to look for an optimal level of human likeness in designing a robot's appearance with consideration of its functions as stipulated by its occupational requirements.

Limitations

There are a number of limitations in the present research. In particular, for Experiment 2, the sample size was small, which may result in the lack of significant findings in that study. Also, the procedures were too artificial, as the participants in the same function condition were told that the four robots shared the same capabilities even though they looked very dissimilar. Future research should use a between-subjects design, and some filler functions should be used to disguise the purpose of the study. For Experiment 3, the limited number of robots used in this experiment restricted us from testing the relation between human likeness and the warmth- and competence-related capabilities and the effect of the relation on the judgment of suitability for different tasks. Future research is needed to further explore this question. Also, although participants were asked to rate their preference for the 29 capabilities for each robot, we are not sure if they continued to assume that the robot possessed the capabilities that they preferred when they rated its suitability for the different occupations. Future experiments should implement a procedure to clarify with the participants that they should assume the robots to possess the capabilities that they preferred.

In conclusion, in three experiments, we found a systematic relation between the appearance of a robot and its perceived capabilities. People infer human- and animal-like robots to be more likely to have warmth-related capabilities. Furthermore, we found that human likeness does not necessarily increase the liking of a robot. However, a robot's potential for emotional capabilities tend to increase people's liking for it. Finally, people judged robots with different capabilities to be suitable for different occupations, preferring robots with emotion capabilities more in occupations that require frequent interactions with humans than in occupations that do not. Robots with increasingly more capabilities and resemblance to humans are capturing the attention and imagination of laypeople and roboticists alike. It is high time that research addresses how the appearance and capabilities of robots can help to advance the presence and acceptance of robots in social and work settings. The present research represents a first step in this direction.

References

Aaker, J., Vohs, K. D., & Mogilner, C. (2010). Non-profits are seen as warm and for-profits as competent: Firm stereotypes matter. *Journal of Consumer Research*, *37*, 277–291.

Bartneck, C., Kanda, T., Mubin, O., & Al Mahmud, A. (2009). Does the design of a robot influence its animacy and perceived intelligence? *International Journal of Social Robotics*, *1*, 195–204.

Carey, B., & Markoff, J. (2010, July 10). Students, meet your new teacher, Mr. Robot. *The New York Times*. Retrieved from http://www.nytimes.com/2010/07/11/science/11robots.html?_r=1&emc=eta1

Christensen, H. I., & Pacchierotti, E. (2005). Embodied social interaction for robots. In *Proceedings of the AISB Symposium on Robot Companions: Hard problems and open challenges in* *human-robot interaction* (pp. 40–45). Hatfield, Herfordshire: The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*, 864–886.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2003)* (pp. 55–60). Los Alamitos, CA: IEEE.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York, NY: Guilford Press.

Kulic, D., & Croft, E. (2005). Anxiety detection during human-robot interaction. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)* (pp. 616–621). Los Alamitos, CA: IEEE.

MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. In *Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward social mechanisms of android science* (pp. 26–29). (Publisher unknown)

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of anxiety toward robots. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2006)* (pp. 372–377). Los Alamitos, CA: IEEE.

Powers, A., & Kiesler, S. (2006). The advisor robot: Tracing people's mental model from a robot's physical attributes. In *Proceedings of the ACM SIGCHI/SIGART 2nd Conference on Human Robot Interaction 2006* (pp. 218–225). New York, NY: ACM Press.

Roesch, S. (2010). Multivariate analysis of variance (MANOVA). Retrieved from http://www.psychology.sdsu.edu/new-web/FacultyLabs/Roesch/Lecture4_MANOVA.ppt.

Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: A psychological approach. *Universal Access in the Information Society*, *4*, 146–155.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Pearson Education.

Torrey, C., Powers, A., Marge, M., Fussell, S. R., & Kiesler, S. (2006). Effects of adaptive robot dialogue on information exchange and social relations. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human Robot Interaction 2006* (pp. 126–133). London, UK: ACM Press.

U.S. Bureau of Labor Statistics. (1999). *Revising the Standard Occupational Classification System* (Report 929). Retrieved from http://www.bls.gov/soc/socrpt929.pdf

Young, J. E., Hawkins, R., Sharlin, E., & Igarashi, T. (2009). Toward acceptable domestic robots: Applying insights from social psychology. *International Journal on Social Robotics*, *1*, 95–108.

Sau-lai Lee is currently an assistant professor at Nanyang Technological University, Singapore. She had previously worked as a post-doctoral fellow at Carnegie Mellon

University for two years. Her research focuses on communication, its psychological processes, and consequence. She is also interested in how human perceive robots, especially how to predict individuals' attitude to different kinds of robots.

Ivy Yee-man Lau is currently a practice assistant professor at Singapore Management University. She received her PhD from the University of Alberta, Canada. Before joining SMU, she was with the Psychology Department at the University of Hong Kong. Her research interests include communication, cultural psychology, and intergroup perception.

Ying-yi Hong is a professor at the Nanyang Business School of Nanyang Technological University, Singapore. She received her PhD from Columbia University in 1994, specializing in experimental social psychology. She taught at the Hong Kong University of Science and Technology from 1994 to 2002 before moving to the University of Illinois at Urbana-Champaign, where she taught at the Department of Psychology for six years. Her main research interests include culture and cognition, self, identity, and intergroup relations.