



Published in final edited form as:

Int J Soc Robot. 2015 April 1; 7(2): 309–331. doi:10.1007/s12369-014-0269-4.

Why Some Humanoid Faces Are Perceived More Positively Than Others: Effects of Human-Likeness and Task

Akanksha Prakash and Wendy A. Rogers

School of Psychology, Georgia Institute of Technology, 654 Cherry Street NW, Atlanta, GA 30332-0170, USA

Akanksha Prakash: aprakash30@gatech.edu; Wendy A. Rogers: wendy@gatech.edu

Abstract

Ample research in social psychology has highlighted the importance of the human face in human–human interactions. However, there is a less clear understanding of how a humanoid robot's face is perceived by humans. One of the primary goals of this study was to investigate how initial perceptions of robots are influenced by the extent of human-likeness of the robot's face, particularly when the robot is intended to provide assistance with tasks in the home that are traditionally carried out by humans. Moreover, although robots have the potential to help both younger and older adults, there is limited knowledge of whether the two age groups' perceptions differ. In this study, younger ($N = 32$) and older adults ($N = 32$) imagined interacting with a robot in four different task contexts and rated robot faces of varying levels of human-likeness. Participants were also interviewed to assess their reasons for particular preferences. This multi-method approach identified patterns of perceptions across different appearances as well as reasons that influence the formation of such perceptions. Overall, the results indicated that people's perceptions of robot faces vary as a function of robot human-likeness. People tended to over-generalize their understanding of humans to build expectations about a human-looking robot's behavior and capabilities. Additionally, preferences for humanoid robots depended on the task although younger and older adults differed in their preferences for certain humanoid appearances. The results of this study have implications both for advancing theoretical understanding of robot perceptions and for creating and applying guidelines for the design of robots.

Keywords

Androids; Humanoids; Older adults; Robot appearance; Robot acceptance; Uncanny valley theory

1 Introduction

The human face plays an important role in human–human interactions. It serves not only as a marker of identity making it possible to visually distinguish one person from another, but also as a medium via which non-verbal social cues are communicated. From an ecological perspective, such cues have been found to fulfill a two-fold purpose [1]. Firstly, facial cues play an adaptive role in human–human interaction by communicating to the observer the

condition, intention, or need of the observed. For example, a baby's cute face signals its vulnerability and evokes protective instincts in the care-taker [2]. Secondly, meanings of such cues are learned and generalized to decipher new faces in unfamiliar contexts (e.g., in forming impressions about the personality of a stranger with a baby-like face; [3]).

Perceptions of facial appearances often influence people's first impressions of others [1,3,4]. There is also tendency to “over-read” or over-generalize the cues emanating from a face. “Attractiveness halo” is a common phenomenon resulting from such over-generalizations [1]. People with attractive faces are rated more positively on a wide variety of dimensions; for example, they are perceived as more intelligent, healthy, and sociable than people with less attractive faces [5]. Additionally, there is evidence that people not only “read from faces” but also “read into faces” [6]. How a face is perceived is also dependent on other available information about the person. For instance, baby-faced adults are more likely to be agreed with than mature-faced adults when both groups' profile descriptions lack in trustworthiness; however, maturefaced adults invoke more agreement than baby-faced adults when both parties are described as less competent, that is, when expertise is questionable [7].

Initial impressions are not always accurate but they affect the behavior adopted. For instance, people are more likely to date a person who is perceived to be physically more attractive [8,9]. Although most of the literature on the social psychology of face perception has focused on attractiveness and baby-facedness, there are other influences of appearance on perceptions. For example, facial appearances also affect perceptions of competence of political candidates, thereby influencing voting choices [10].

Extrapolating from research on the perception of human faces, people's initial impressions about human-looking robots might involve similar socio-cognitive processes as found in human–human interactions. There is evidence for a human tendency to over-use social rules and expectations and apply them onto machines such as computers [11]. For example, individuals' behaviors towards computers were found to be consistent with the social rules of politeness, reciprocity, and retaliation (see [12] for a detailed review). The generalizations of social rules and expectations might be heightened in situations where human–machine interaction matches human–human interaction. That is, when a robot is designed to look and behave like a human and assist with tasks that are traditionally performed by humans, people may ascribe human-like attributes and expectations onto the robot. This may in turn impact their behavior toward and acceptance of the robot in question.

1.1 A Wide Variety of Humanoids

Conventionally, the use of robots is identified with military, manufacturing, and space-research domains. However, robot applications are now extending to domestic, healthcare, and entertainment settings. Thus, although various service robots are being developed for personal use in the home environment, the end users, in general, cannot be expected to understand the internal architecture and operations of the robotic system. Thus, for successful deployment and adoption of such technology, it should be perceived as easy to use [13].

Many roboticists believe that because people are most accustomed to interacting with humans, giving robots human form and functionality should make human–robot interaction easier and qualitatively better [14]. This is at least one of the reasons for the increased zeal in the design of “humanoids”, which are robots designed to have some human resemblance. Extreme resemblance to a human marks the category of “android robots”. Thus, androids are humanoid robots that are designed to be almost indistinguishable from a human in appearance and behavior [15].

Despite the prevalence of humanoid robots, the appearance of humanoids varies to a great extent. Almost all humanoid robots have heads, but there are variations in the head shapes (e.g., round vs. rectangular, wide vs. narrow) and also in the composition of facial features. Therefore, researchers are trying to develop criteria that can enable the assessment of a robot's human-likeness. A study on the design of humanoid heads found that 62 % of the variation in the perception of human-likeness of humanoid heads was accounted for by the presence of particular facial features [16]. Nose, eyelids, and mouth were found to be the facial features that provided the most enhancements to a robot's human-likeness. The study also found that if the head was wider than it was tall, the robot was perceived as more robotlike (and therefore, less human-like) in appearance. Moreover, the appearance became less human-like if the proportion of head-space for forehead, hair, or chin was reduced.

Such criteria can be used to evaluate the human-likeness of a robot's face. However, how people perceive robots that differ in human-likeness is not fully understood. Although attempts have been made to unveil the effects of humanlikeness, the findings in the existing literature are inconsistent. Thus, one of the primary goals of this research was to understand people's perceptions of robot faces across a range of human-likeness in appearance.

1.2 The Uncanny Valley Theory

Studies on the perception of humanoid robots are often aimed at investigating the validity of the uncanny valley theory [17]. The uncanny valley theory is probably the most popular theory on the effect of robot appearance. The theory relates the human-likeness of a robot with the level of familiarity (affinity in the later translation [18]) evoked in the human observer. This relation is posited to be curvilinear. According to the theory, as the appearance of a robot increases in human-likeness, people's familiarity with it increases. However, this relationship presumably holds only up to a certain point. Beyond this critical point, further increase in robot human-likeness reduces the level of familiarity, that is, the robot is then perceived as strange or eerie. If the robot's human-likeness can be increased to almost entirely match the appearance of a human, familiarity will rise again and will be maximized when the robot cannot be distinguished from a healthy person. The region of dip in familiarity (where the robot is perceived as strange) with increasing human-likeness is referred to as the uncanny valley (see Fig. 1).

The main limitation of the uncanny valley theory is that it was formulated based on anecdotal examples. At its conception, the theory was not experimentally verified, but was proposed by Mori as a generalization of his experiences with prosthetic hands, mannequins, and robots [15]. However, despite its non-empirical basis, it has triggered a plentitude of

research, probably due to its historical significance as one of the earliest theories of perceptions of human-like robots (e.g., [15,19–24]).

To assess if the uncanny valley can be plotted as Mori hypothesized, MacDorman and Ishiguro morphed robot faces onto human faces in different proportions and used them as stimuli to gauge people's perceptions [15]. The study was conducted with Indonesian participants through computer-based questionnaires. Participants saw the still images of the generated faces and rated them on scales for human-likeness (1=very mechanical; 9=very human-like), familiarity (1 =strange; 9=very familiar), and eeriness (0 =not eerie; 10 =extremely eerie). People's familiarity ratings when plotted against perceived robot human-likeness ratings resulted in the uncanny valley pattern.

In a later study, a different set of Indonesian participants viewed *video clips* of 13 existing robots and a human, and rated each robot and the human on human-likeness, familiarity, and eeriness [25]. In this case, the plots of human-likeness versus familiarity did not reveal the U-shaped plot predicted by uncanny valley theory. Due to such discordant findings with the varying nature of stimuli, it remains unclear what factors are most influential in the initial perception of humanoid robot faces.

Although still images have limitations of not displaying subtle, dynamic movements and expressions that may further impact perceptions of the beholder, videos are also not without constraints. The apparent caveat in the MacDorman [25] study was the use of a wide range of robots performing dissimilar actions in different settings. It is, therefore, not possible to decipher how participants' appraisal of the robots' activities and environments informed their impressions of the robots' appearances. Moreover, an added confound in the study's design was that some of the robots had voices and others did not. Thus, due to lack of systematic manipulation of relevant variables, participants' ratings of the videos did not elucidate which of the robots' features and actions were being attended to and were influencing the formation of perceptions.

In both of these studies, the participants viewed the robots and rated them without any (actual or imagined) context of interaction with the robots. It is plausible that if the viewers were to perceive the robots as performing a relevant task for them, their ratings of familiarity would be more indicative of their attitudinal acceptance of robots. Familiarity in itself is an ambiguous construct and is not informative of people's preferences. High familiarity may not necessarily imply liking or acceptance. Similarly, low familiarity may not always imply disliking or rejection. Probably, because the translation of Mori's original article [17] used the term “familiarity”, researchers investigating the uncanny valley continued to measure how familiar people find robots based on their appearances. However, more specific measures need to be used to gain clearer insights into people's perceptions of humanoid robots.

The uncanny valley theory's assumption that human-likeness and machine-likeness are the extreme ends of the same dimension has been questioned in recent studies [26]. Certain robot pictures were rated as low on human-likeness but also less “mechanical” suggesting that these need to be considered as separate dimensions in the assessment of appearance

[26]. Moreover, the criteria on which participants evaluate human likeness of a robot appearance are not explicit in surveys. Therefore, controlled experimental studies are needed that can show the effects of varying levels of human-likeness on participants' perceptions. For example, Burleigh and colleagues carefully manipulated human-likeness on digital faces by varying the sizes of different facial features and by merging human and non-human features to various degrees, and found that a face was perceived as eerie when its category membership was not clear (e.g., when human and non-human features were almost equally merged [27]).

1.3 Perceptions in the Context of Robot Task

The home setting is increasingly being considered as a potential market for service robot applications. A wide range of robots of varied appearances are currently under development that may potentially assist with everyday living tasks (for a review, see [28]). Such assistive robots have the potential to support people's independence and well-being, and can be specifically beneficial in helping people age successfully in their own homes.

There are many tasks that people must perform to maintain their independence and health, including self-maintenance, instrumental, and enhanced activities of daily living [29,30]. Self-maintenance activities of daily living (ADLs) include the ability to toilet, feed, dress, groom, bathe, and ambulate. Instrumental activities of daily living (IADLs) include the ability to successfully use the telephone, shop, prepare food, do the housekeeping and laundry, manage medications and finances, and use transportation. Enhanced activities of daily living (EADLs) include participation in social and enriching activities, such as learning new skills and engaging in hobbies. Age-related changes in physical, perceptual, and cognitive abilities may make performing these tasks more challenging for older adults [31]. Even for younger people, robots may play a beneficial role by saving them time and effort. Moreover, high workload in people's professional lives may prevent them from regularly taking care of some household activities. Well-designed, functional robots can facilitate timely maintenance of the home and its surroundings, provide entertainment and companionship, help in solving cognitively challenging, intellectual problems, and also assist in personal care tasks, if needed.

People's attitudes toward robot assistance vary with the task [32–34]. Older adults are selective in their preferences of robot assistance over human assistance [34]. The selectivity is determined by the nature of the home-based task. In general, there is higher attitudinal acceptance of robots for assistance with IADLs (e.g., chores), followed by EADLs (e.g., learning a new skill). Older adults were least open to robot assistance for ADLs (e.g., bathing). However, these findings were likely influenced by the specific type of *robot* being considered. For example, a robot's appearance creates expectations for what it can or cannot do successfully and may in turn affect people's openness to take robotic assistance for certain tasks [35].

Thus, although task seems to determine the overall acceptance of a robot, less is known about the effect robot task will have on how people react to *human-likeness* in robot appearance. Are there tasks for which a highly human-like robot would be evaluated more positively than less human-like appearances? Are there tasks for which the trend would

reverse? In sum, the open question is “how do task and human-likeness jointly impact perceptions of robots?”

It has been suggested that an appropriate match between a robot's appearance and its task can improve people's acceptance of the robot [36]. Goetz et al. found that people are likely to prefer human-looking robots to perform jobs that entail more social skills (e.g., sales representative, aerobics instructor) but greater preference would be shown for machine-looking robots for jobs less social in nature (e.g., customs inspector, security guard).

It is worth noting that the human-like robot faces used in the Goetz et al. study [36] were not fine imitations of human faces. The stimuli used as human-looking faces can be considered more human-looking than the machine-looking faces employed in the study. Nonetheless, they did not resemble the appearance of a person in that the faces were simplistic, cartoon-like renditions of human facial shape, features, and hair, and less sophisticated in details. Thus, though informative about the interactive effect of robot appearance and task, the Goetz et al. study [36] did not provide insights into people's perceptions of very human-looking robots. For instance, if a robot were designed to look indistinguishable from a human, which tasks would it be most preferred for?

Robot appearance and task also had an interaction effect when the robot played the role of a co-worker in a work environment [37]. People felt more responsible when working with a machine-looking robot than when working with a human-looking robot, particularly when the robot was in a subordinate position. Based on this finding, Hinds et al. suggested that robots should be made mechanical-looking when assisting in environments where personal responsibility is important [37]. However, the researchers used two extreme manipulations of robot appearance such that the mechanical looking robot did not have a human form whereas the human-looking robot looked like a white male. Thus, their study did not unveil the impact of intermediate human–robot appearance; that is, how would a robot with mixed human–robot features be perceived? Moreover, how would perceptions of such a robot be influenced by its task?

1.4 Age-Related Differences in the Perception of Robots

The plethora of research on the uncanny valley theory notwithstanding, there is lack of concrete evidence to infer the effects of a robot's human-likeness on the human perceiver. Some major methodological limitations of the research on uncanny valley were discussed in the previous section. Another primary limitation is the narrow range of potential users; most research on robot appearance has involved only young adult participants. To test the validity of any general theory on robot perception, a broader age-range of participants should be considered.

Considering different age groups is essential not only to enhance a theoretical understanding of how differences in abilities and experiences affect robot perception but also to improve the design of robots so they are more acceptable to the target users. For example, age-related differences might lead to differences in needs and preferences. From the perspective of application, robots have the potential to support people's independence and well-being. They can specifically assist older adults with various home-based tasks, so they can continue

living independently in their homes [38]. The limited human–robot interaction (HRI) studies conducted with older adults have shown that although older adults have less experience with robots [33], they have expectations and opinions about robot appearance, particularly in terms of size [39,40]. However, we have scarce knowledge of how older adults perceive highly human-looking robots (androids) in comparison to less human-looking ones (humanoids).

There are also gaps in our understanding of younger adults' perceptions of humanoid robots in comparison to androids. Researchers often focus on comparing perceptions toward mechanical appearance (devoid of any human features) with humanoid robots. For example, a study conducted with university undergraduates concluded that younger adults in general showed preference for human-like appearance of robots although large individual differences in preferences were noted [41]. However, even the most human-like appearance manipulated in the study had some human features but was not close to a human appearance. It remains unclear if and how younger adults' perceptions would change if the appearance were further increased in human-likeness to match the appearance of a person.

Although more research has been conducted with younger adults than with older people, we have limited knowledge of both the age groups' perceptions of humanoid and android appearances for robots. Understanding perceptions of a broad range of users can guide the design of robots that are acceptable to the target user.

1.5 Measuring Perceptions

Determining appropriate and informative measures of participants' reactions and attitudes towards robot appearance is another challenge in robot appearance studies. Not only do we need to select suitable outcome variables with clear definitions, but also establish the best methodology to deploy the measures (e.g., through questionnaires, experiments, and/or interview studies).

1.5.1 What Do “Perceptions” Comprise?—Initial perceptions of a robot can be based on different kinds of appraisals people make, such as, how useful it is perceived to be, how much trust it evokes, how likeable it seems, and how anxious it makes them feel. These appraisals seem related but the strengths of the inter-correlations are unknown. People are also likely to ascribe relative importance to every factor. For example, if a robot is considered useful even if it is not liked, which factor will have a greater impact on overall perception of the robot?

Common measures used in studies investigating the uncanny valley theory include: affect evoked such as fear and anxiety [42]; attractiveness versus repulsiveness [19,43]; familiarity [15,25]; likeability [21,43–45]; and perceived eeriness [15,25]. Each of these measures informs about a particular constituent of perceptions; however they cannot independently provide a complete picture of perception-formation. The need is to evaluate perceptions on multiple dimensions together for a holistic understanding of *attitudinal* acceptance of robots. Attitudinal acceptance, defined as the users' positive evaluation or beliefs about the technology, has been argued for as a precursor to behavioral acceptance, that is, the users' actions in using the product or technology [13].

1.5.2 Variables Assessed in Technology Acceptance Models—Robots are advanced technology and the general technology acceptance models can be a starting point for the conceptualization of robot acceptance. The technology acceptance model (TAM; [13]) is the most widely recognized model of technology acceptance. The TAM was developed to understand prospective expectations about information technology usage. The model includes two main variables that affect acceptance: perceived usefulness and perceived ease of use. There is strong empirical support for the TAM [46,47], in part due to its ease of application to a variety of domains. However, the model's simplicity has evoked some criticism [48] that has led to the development of other models. For example, the unified theory of acceptance and use of technology (UTAUT) model was developed with the intent of unifying a large number of acceptance models [47]. UTAUT posits that technology acceptance may be determined by the following constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions. In an alternative model, the technology-to-performance chain model (TPC; [49]), technology adoption is presumed to be impacted by the technology's utility and its fit with the tasks it is designed to support (referred to as task-technology fit).

Robots differ from other technologies in certain aspects, and the existing technology acceptance models may not provide a complete picture for robot acceptance if applied without modifications. For instance, a personal robot is an *embodied* agent with social capabilities and social presence, and the expectation for it is to work in a collaborative manner with the user [50]. This may heighten the importance of its appearance or human-likeness, the characteristics of tasks it performs, and the affect it evokes in the user. TAM and UTAUT do not include variables of appearance, task characteristics, and affect. The TPC model is oriented toward information technology and, even with a task-technology fit dimension, may not be suitable for an embodied agent that has a social presence outside of the computer system.

It is also worth considering that robots have been a topic of science fiction literature and film for decades. Rosie from the Jetsons, C3P0 and R2D2 from Star Wars, and Robby the Robot from Forbidden Planet, are all well-known science fiction characters that may have influenced the way in which the general public thinks about robotics. Likewise, fictional robots in antagonistic roles such as the Terminator propagate a negative image of robots. Thus, media exposure may create preconceived expectations about robots, even for individuals who have never interacted with a robot directly. In fact, people do have ideas or definitions of what a robot should be like [39]. Pre-existing ideas about robots may lead to evaluations of an existing robot against criteria based on one's expectations. Violations of expectations are likely to negatively impact acceptance.

1.5.3 Robot Attitude Scales—Psychological scales have been developed to measure people's perceptions of robots, with the most widely recognized scales being the negative attitude towards robots scale (NARS; [42,51]) and the Robot Anxiety Scale (RAS; [52]), which are used to gauge psychological reactions evoked in humans by robots. These scales assess to what extent people feel unwilling to interact with a robot due to arousal of negative emotions or anxiety.

The NARS assesses negative attitudes toward robots considering three dimensions: interaction with robots, social influence of robots, and emotional interactions with robots. The RAS also has three dimensions or sub-scales: anxiety toward communication capability of robots, anxiety toward behavioral capability of robots, and anxiety toward discourse with robots. It can be used to assess state-anxiety in real or imaginary interactions with robots. The limitation of both the NARS and RAS scales is that they focus only on negative affect and lack measures of positive evaluations of the robot and interactions with it. Moreover, the scales do not provide any understanding of the underlying causes of negative affect toward robots. For instance, anxiety toward a robot may result from participants' mental-models or stereotypes against robots. However, anxiety can also be triggered due to lack of familiarity with robots in general, and thus can be eased over time as participants become more accustomed to the robots.

More recently designed scales are oriented toward both negative and positive attitudes toward robots. The robot attitude scale (also abbreviated as RAS; [53]) is one such scale in which a robot is rated from 1 to 8 on 11 dimensions: safe–dangerous, reliable–unreliable, friendly–unfriendly, simple–complicated, useful–useless, strong–fragile, interesting–boring, trustworthy–untrustworthy, advanced–basic, easy to use–hard to use, and helpful–unhelpful. Similarly, the Almere model, an adaptation of the UTAUT, is aimed at understanding older adults' acceptance of assistive social robots and has nine constructs: anxiety, attitude towards technology, facilitating conditions, intention to use, perceived adaptiveness, perceived enjoyment, perceived ease of use, perceived sociability, perceived usefulness, social influence, social presence, trust, and use [54]. These scales and models are useful developments in the space of human–robot interaction. However, their purpose is limited to identifying general trends without in-depth explanations for *why* people hold certain perceptions.

1.6 Overview of Present Research

The current literature on perception of humanoid robots has identified important variables such as robot's appearance, task, and user characteristics that can affect perceptions of robots. However, in most cases, these variables are defined and manipulated differently and are often studied in isolation from other variables. Thus, the information gauged from such studies is difficult to integrate into a holistic understanding of people's initial perceptions of robots. This study was designed to address this gap in the existing research by exploring:

1. How do people's perceptions of robot faces vary for a range of human-likeness in facial appearance?
2. Do perceptions of robots of different levels of human-likeness vary across tasks?
3. Do younger and older adults differ in their perceptions of robots?

Four tasks were selected and individuals were instructed to imagine the robot assisting them with the completion of each task. These tasks exemplified each of the three categories of daily living activities (i.e., ADL, IADL, and EADL). The ADL and EADL categories were represented by one task each. The IADL category was instantiated through two examples to represent the level of cognitive demand. Activities such as chores in the home are IADLs

with low level of cognitive demand whereas finance and medication management can impose high cognitive load on the individual.

Four dependent variables were used to assess people's perceptions of robots: likeability, anxiety, trust, and perceived usefulness, in the four task contexts. These variables were selected to represent the range of variables assessed in the literature, and capture both affective (i.e., likeability, anxiety) and cognitive (i.e., perceived usefulness) components of individuals' attitudes (affective events theory; [55]). Trust incorporates both affective and cognitive components [56,57]. Likeability can indicate whether people would generally like a robot that has a certain level of human-likeness to assist them with a particular task. Trust in a robot is a predictor of its acceptance and is a dimension used in the most recent robot attitude scales [53,54]. Perceived usefulness was measured as it is one of the main variables in the technology acceptance model [13]. Anxiety is frequently used in the assessment of human–robot interaction. For example, the robot anxiety scale (RAS; [52]) is solely focused on the measurement of anxiety. The more recently developed Almere model also includes a measure of anxiety [54]. Thus, the goal behind using multiple dependent variables was to gain a holistic understanding of people's perceptions of robots.

Participants imagined being assisted by robots in the aforementioned task contexts. For every task, they rated different robot pictures shown on a computer screen. At the end of this rating task, participants were briefly interviewed about their preferences for one robot face over the others. Additionally participants completed various questionnaires at different points in the study. The goal behind using a combination of these methods was to assess the trend in people's reactions to robot faces in different task contexts and to understand the underlying reasons for the same. A part of this study was reported in Prakash and Rogers [58].

2 Method

2.1 Participants

The participants were 32 younger adults (18 females) and 32 older adults (19 females). The younger adults ranged in age from 18 to 23 ($M = 20.16$, $SD = 1.42$); the older adults were between the ages of 65 and 75 ($M = 70.09$, $SD = 3.07$). All younger adult participants were recruited from the Georgia Institute of Technology undergraduate population, and received credit for participation as a course requirement. The older adults were recruited from the Human Factors and Aging Laboratory database and were compensated \$36 for their participation. A majority (78 %) of older adults reported having a college or higher degree.

The samples were diverse in race/ethnicity. In the younger adult group, 62.5 % reported themselves as White/Caucasian, 6.25 % as Black/African American, and 28.12 % as Asian. One younger adult did not report her race. In the older adult group, 62.5 % reported themselves as White/Caucasian, 34.37 % as Black/African American, and 3.12 % as multi-racial.

Participants also provided general information about their health. They responded to the questionnaire item “In general, would you say your health is...” on a five-point scale

(1=poor; 3=good, 5 =excellent). On average, both younger adults ($M = 3.81$; $SD = 0.86$) and older adults ($M = 3.66$; $SD = 0.94$) reported having good health.

2.2 Materials

We report here only the materials that are relevant for the focus of this paper. Additional questionnaires were also used as part of the larger study [59].

2.2.1 Humanoid Pictures—To manipulate levels of human-likeness, four robot faces and four human (two male, two female) faces were selected. The human faces were chosen from the Montreal set of facial displays of emotions. All four faces had neutral expressions and were of White Caucasian individuals of similar ages. The four robot faces corresponded to the humanoid robots: Pearl Nursebot, Nexi MDS (Mobile/Dextrous/Social), Nao, and Kobian. Humanoid robots vary in appearance characteristics in an unsystematic manner. Therefore, robot stimuli were chosen for this study to cover a range of humanoid facial appearance (Fig. 2). For example, there were variations in the shapes of the faces. Whereas Nexi and Pearl are somewhat round-faced, Nao and Kobian appear to have rectangular faces. Nao's face is wider than long and the opposite is true for Kobian. Moreover, differences in the number of facial features were also evident. Nao's face is minimalistic in design; other three faces have more features but of different sizes and shapes. Pearl and Kobian seemed to have ear-like structures as well which Nexi and Nao lack. Moreover, Nexi and Kobian seem bald but Pearl and Nao have impressions of hairlines. However, all four robots had a pair of eyes and a resemblance of a mouth and/or a nose.

Each robot face was paired with a human face. For each robot–human pair, we created an appearance that lay between that of the human and the robot by morphing the two pictures. Thus, for every robot-human pair, the participants would see three appearances: completely human-like, between human-like and robot-like, and completely robot-like. In all, four such sets of face pictures were generated resulting in a total of 12 pictures (Fig. 2). All pictures were converted to black and white and were cropped to be of the same size. The pictures were printed on separated sheets and laminated to be presented to the participants during the interview.

2.2.2 Questionnaires—We collected data for four questionnaires:

1. Demographics questionnaire [60] was administered to collect demographics and general health information.
2. Robot opinions questionnaire [34] is a robot-specific 12-item questionnaire modeled after technology acceptance scales [13]. Participants respond to six questions on perceived usefulness, such as “I would find a robot useful in my daily life”, “Using a robot would make my daily life easier” and six questions on perceived ease of use, such as “My interaction with a robot would be clear and understandable” on a seven-point Likert scale (1=extremely unlikely, 4=neither unlikely or likely, 7=extremely likely).
3. Robot facial appearance questionnaire assessed people's opinions about the facial appearance of their imaginary robot and was developed for this study. It consists of

15 items (e.g., “I would want my robot to have eyes”, “I would want my robot to look exactly like a human”, and “I would want my robot's face to be unique”). Responses are marked on a five-point Likert scale (1=strongly disagree, 3=neither agree nor disagree, 5 =strongly agree).

4. Robot familiarity and use questionnaire (administered in previous studies such as [33,34]) was a 13 item questionnaire that required participants to indicate their level of experience with 13 different kinds of robots on a five-point scale (0 =not sure what it is; 1=never heard about, seen or used this robot; 2=have only heard about or seen this robot; 3 =have used or operated this robot only occasionally; 4 =have used or operated this frequently).

2.3 Design

The rating task was a 2 (age) \times 3 (human-likeness) \times 4 (task) split plot design where age group was a between subjects factor; human-likeness and task were within subjects factors. Age group consisted of two levels: younger adults and older adults. Human-likeness was the degree to which the robot face resembled a human face and comprised three levels: human appearance, mixed appearance, and robot appearance (Fig. 2). Task had four levels: personal care (ADL), chores (IADL; low cognitive demand), decision-making (IADL; high cognitive demand), and social task (EADL).

Participants' perceptions were assessed via four dependent measures (DVs): perceived usefulness, trust, likeability, and anxiety. Each DV consisted of a single item and the response was measured on a five-point scale where 1=not at all, 3=a fair amount, and 5=very much.

2.4 Procedure

Participants signed an informed consent followed by the demographics questionnaire and the robot opinions questionnaire. Next, they performed a rating task on a computer.

2.4.1 Rating Task—Participants were first given an overview of the rating task followed by a practice task. The first part of the practice task consisted of getting familiarized with the keys for selecting one's responses. During the second part of the practice, participants were given a sample rating task wherein they were asked to imagine that they are at a department store and a robot was assisting them in finding items of their choice. Participants were then shown pictures of robots (different from the ones used in the study) and were asked to rate those on perceived usefulness, trust, likeability, and anxiety. At the end of practice, participants' questions, if any, were answered. Participants were also informed that the speed at which they responded did not matter and that there were no right or wrong answers.

When the participant was ready to begin the rating task, the following set of instructions was displayed on the screen and was also read aloud (to minimize differences in participants' assumptions about robot capability, autonomy, and control):

“Imagine that you need some assistance and that you have been given a robot to take home with you.

- The robot can perform tasks for you.
- You do not have to program the robot.
- You should assume that the robot can do what you want it to do.
- In this study we are focusing on the robot's face. Assume the robot's body to be consistent with the robot's facial appearance. The robot's body is such that it does not reduce its efficiency in performing a task.”

Next, participants were asked to imagine interacting with a robot in four different task scenarios (one scenario at a time with breaks in between the scenarios). The details of the task scenarios are delineated in Fig. 3. In each scenario, participants were presented with 12 face pictures (human, mixed, and robot; four of each; Fig. 2). They were asked to imagine the robot to have the appearance as shown in every picture and then rate the robot in terms of how useful they would find it, how much they would trust it, how much they would like it, how anxious they would feel toward it in that task scenario. Thus, even when participants saw a human face, they were instructed to imagine it to be the face of a robot. The participants provided their responses on a five-point unipolar Likert-type scale where 1 = “not at all”, 2 = “a little”, 3 = “somewhat”, 4 = “much”, and 5 = “very much”. The presentation of the pictures was randomized without any constraints. Task contexts and rating measures were counterbalanced using a 4×4 partial Latin square design. Thus, in every task context, each picture was rated four times (once for each of the four DVs). Every participant completed a total of 192 ratings (=12 pictures \times 4 DVs \times 4 tasks).

2.4.2 Appearance Preference Interview and Questionnaires—At the end of the rating task, participants were taken to another room where they were interviewed about their preferences for the robot facial appearance. The interview was audio-recorded. Participants were asked to imagine that they own a robot that stays with them in their home. The robot can assist them with all the tasks in the home that they imagined earlier during the rating task (i.e., it can bathe the person, perform daily chores, help in making investment decisions, and also provide them social companionship.)

Participants were presented with the two female human pictures (Fig. 2) and were asked, “Which of these two faces would you prefer your robot to have?” The same was repeated with the male human pictures. The preferred female picture was then placed adjacent to the preferred male picture and participants were asked to decide which one they would prefer over the other. They were also asked to provide reasons for their selection. Similar selection tasks were performed for the mixed and the robotic appearance pictures.

Finally, the participant's most preferred human picture, the most preferred robot picture, and the most preferred mixed picture were placed together in front of the participant. The participant was asked to pick the most preferred appearance out of the three pictures, and also provide reasons for the choice. Next, participants were asked to think specifically about a personal care task (e.g., bathing) and discussed if they would have a preference among the three faces if the robot helped them with that task. This was repeated for the other three categories of tasks: menial task (e.g., chores), social task (e.g., chatting with someone, playing a game with someone, or learning a new skill from someone), and decision-making

task (e.g., deciding where to invest money). The order of the tasks was held constant for the interview. The rationale was that the chances of order effects were much less at this point as participants would have had enough exposure to every picture and to the four task contexts.

After the interview, participants completed the robot facial appearance questionnaire, the robot familiarity and use questionnaire, and the assistance preference checklist. At the conclusion of the study, participants were debriefed and compensated for their time.

3 Results

3.1 Questionnaire Analysis

The results from the quantitative analysis of the questionnaires are presented first to provide a description of participants' general attitudes towards robot, their familiarity with and use of existing robots, and their opinions about different features and characteristics that a robot's face should or should not have. In the next section, we focus more specifically on the effects of robot human-likeness by presenting the results from the quantitative analysis of the rating task and the qualitative analysis of the robot preference interview.

3.1.1 Attitudinal Acceptance, Robot Familiarity and Use—The younger and older adults had similar generally positive opinions about robots as assessed through the perceived usefulness and perceived ease of use scales of the robot opinions questionnaire (see Table 1). Both the scales had high internal consistency (greater than 0.90) as measured by Cronbach's alpha.

With respect to familiarity and use, younger adults had more familiarity overall. However, for most of the robots listed, participants in both age groups indicated that they either had no idea of the robot in question or had only heard of or seen it, but did not have experience using it (see Table 1 for mean familiarity ratings for each type of robot).

3.1.2 Desired Characteristics in a Robot's Face—The facial appearance questionnaire, consisting of 15 items, was designed for this study to assess the facial features and facial characteristics that people want their robot to have in general. Wilcoxon signed rank test was performed (separately on the two age groups) to assess if younger and older adults differed from the neutral point of three (i.e., the “neither agree nor disagree”) and in which direction. Wilcoxon rank-sum test was performed on each of the fifteen items to compare younger and older adults' responses. The test statistics and medians are presented in Table 2.

In general, both the age-groups wanted their robot to have a face, eyes, and a mouth. Nose and ears were features that were less strongly desired. In terms of other characteristics, both the age-groups wanted the robot's face to be round in shape, unique, attractive, and expressive. Both groups had neutral opinions for the robot to have male features. However, younger adults desired female features in their robot but older adults were neutral on that item as well.

Significant age differences were observed with regard to some features. Compared to younger adults, older adults were more likely to want their robot to have lips and soft

skin. For both these features, older adults' responses were significantly above the neutral point whereas younger adults neither agreed nor disagreed.

In comparison to younger adults, older adults were also more likely to want their robot to have hair on the head. However, for this feature, the one-sample Wilcoxon test was significant for younger adults in the negative direction, implying that younger adults *did not* want their robot to have hair on its head. In contrast, the older adults responded in the positive (agreement) direction although not significantly above the neutral point.

Younger and older adults' responses were also significantly different on the item "I would want my robot to look exactly like a human" (Fig. 4). Whereas older adults agreed that they would want a robot that looked exactly like a human, younger adults tended to disagree although their median was not significantly different from the neutral point.

3.2 Analysis of Rating Task and Interviews

The results from the rating task and the robot preference interview are presented together. The goal is to provide a holistic picture of the quantitative patterns observed in the perceptions of varying levels of human-likeness in robots, generally and across different tasks, in conjunction with the underlying reasons for the trends.

3.2.1 Quantitative Analysis of Rating Task—There were four dependent variables (DVs): perceived usefulness (PU), trust, likeability, and anxiety. The anxiety data are not included in the main analysis because the term "anxious" was not clearly understood by all participants and some participants were confused with the direction of the scale. For more details on problems with using anxiety as a construct and its analysis see [59].

Separate univariate ANOVAs were conducted on each of the three DVs. The type I error rate was Bonferroni corrected (0.05/3). Therefore, the critical alpha level was set at $p < 0.0167$ for all further analyses. Huynh–Feldt corrections were applied where sphericity assumptions were violated.

The three DVs were highly correlated, especially for older adults (see Table 3). Moreover, trust and PU were more strongly correlated than likeability and trust, and likeability and PU. Such a high correlation between trust and PU implies that the two DVs might be measuring the same underlying construct. As such, we focus here on likeability data and report trust and PU only where differences were noted in the results.

3.2.2 Qualitative Thematic Analysis—To understand the reasons for participants' preferences for one appearance of robot over the others, the audio recordings of the 64 interviews were transcribed verbatim. The primary researcher developed a coding scheme based on the extant literature on robot appearance and social psychology (Table 4). The coding scheme categorized the reasons participants gave for selecting a particular robot appearance. If participants' reasons did not fit into any of the categories of the coding scheme, the scheme was modified to be inclusive of the new response. The primary coder and a secondary coder coded the same two transcripts using MAXQDA text analysis

software and were in 100 % agreement. Thereafter, the remaining interviews were analyzed only by the primary coder.

3.3 How do Perceptions Vary for a Range of Human-Likeness?

Participants' likeability ratings depended on the human-likeness of the robot face. The main-effect of human-likeness on likeability was significant ($F(1.48, 91.97) = 7.32, p = 0.003; \eta_p^2 = 0.10$). Post-hoc comparisons revealed that mixed appearance was liked less than human appearance ($t(63) = -2.85, p = 0.01, r = 0.34$) and robot appearance ($t(63) = -4.75, p < 0.001, r = 0.51$; Fig. 5). Thus, the likeability data offer support for an uncanny valley pattern in that perceptions were most positive for less human-like and completely human-like appearance compared to an appearance that was in between the two levels of human-likeness. Similar trends were noted in trust ($F(1.43, 88.54) = 4.42, p = 0.02, \eta_p^2 = 0.07$) and PU ratings ($F(22.56, 336.59) = 4.16, p = 0.03, \eta_p^2 = 0.06$) but the effects were not statistically significant at $p < 0.0167$.

During the interview, participants were asked to select their most preferred appearance for their robot. The frequency distribution of preferences is illustrated in Fig. 6. An uncanny valley trend was again evident particularly for the older adults with a lower preference of the mixed appearance over the completely robotic or human appearance. A Chi-square analysis assessing whether the distribution of the most preferred face across human, mixed, and robot appearance categories depended on age-group was significant ($\chi^2(2, N = 64) = 8.02, p < 0.05$). Chi-square analyses were also conducted to assess if human, mixed and robot appearances would be preferred by an equal number of people. This analysis was conducted separately for older and younger adults. Older adults' preferences were on the extremes; that is, they were in favor of a completely human-looking appearance (56 %) or a robotic appearance for their robot (37 %), but would not prefer a mixed appearance ($\chi^2(2, N = 32) = 12.25, p < 0.05$). However, younger adults' preferences were more varied; half of them preferred robotic appearance, a fourth preferred mixed appearance and another fourth selected human appearance ($\chi^2(2, N = 32) = 4.00; p > 0.05$).

3.3.1 Reasons for Preferring a Human Face—Fifty-six percent of the older adults and 25 % of the younger adults selected a human face as their most preferred appearance for their home robot. Their major reason for their selection was the degree of human-likeness of the face. Many participants elaborated this further and mentioned that a human-like face would be more familiar and “relatable” than other robotic appearances, for example, “I guess that whole idea of having a robot kind of freaks me out a little bit. Um, so yeah. I like that it looks like a human. I feel like I could connect better with it.” Moreover, people also considered a human-like appearance more apt for fulfilling companionship needs, as is exemplified in this remark from an older adult, “It's not only capable of doing chores functionally but it has within it the capability of being a companion. So your companions look more like you or resemble you, something that's familiar and that one does it.”

Some participants perceived a human-like robot to be more capable in general than other robots, for instance, “...she's more able to perform the duties she's supposed to, I guess.” In fact, a few participants considered human-looking robots to be “the most developed” kind of

robots. In some cases, preferences were influenced by the robot gender as well: “Uh it's a lady, um and that would be just a good companion or somebody for me to talk to, work with...”. On the contrary, another participant who favored a male human appearance said, “And I know you said each one can do everything, but this one kind of is uh, probably preconceived idea that the man-looking robot would be able to do everything.” The perceived personality or expressiveness of the human face was also pointed out as reasons for preference as noted in descriptors such as “smart”, “caring”, “non-intimidating”, and “ [has] a little smile”.

3.3.2 Reasons for Preferring a Robotic Face—Nearly 40 % of the older adults and 50 % of the younger adults selected a robotic face as their most preferred appearance for their robot. Most of these participants did not want their robot to resemble a human because it would be difficult for them to distinguish such a robot from a human: “Well, when I think of a robot, I think of him not being people. Umm these look very realistic and I might confuse a person with a robot. This [the selected picture] is definitely a robot and this I would be in command of.” People also preferred robotic appearance due to the perceived personality or expressiveness of the face (e.g., they found the appearance to be “cute”, “friendly”, “child-like”, and/or “trustworthy”).

A robotic appearance was also favored by participants who tended to ascribe negative human traits or intentionality onto human-looking robots. For instance, a person reasoned why she would not like a human-looking robot, “She just looks like she could tell me a lie. Just be like Yesma'am, yes ma'am, and in the back of her mind she's like ‘I can't wait to get out of here.’” Such participants perceived robotic appearance robots to be devoid of such flaws: “It seems like it will do exactly what it's supposed to do.”

3.3.3 Reasons for Preferring and Not Preferring a Mixed Appearance—Only about 6 % of the older adults preferred a mixed appearance to human and robotic. However, 25 % of younger adults preferred a mixed appearance. These participants explained that a mixed appearance was better than the extremes because it employed the benefits of human and mechanical appearance. A quotation from a young adult exemplifies this reasoning, “... because although it's human enough to be familiar, it's, like, clearly not human so...I still perceive it as like a robot, but it doesn't make me as uncomfortable as a human face would on a robot.” A similar justification was noted in another participant's comment: “It's not quite as I guess invasive as having another person living with you, but it's not as unrealistic as having like a robot from a horror movie or something living with you. It's a good blend of both”.

Many participants, particularly older adults, who did not favor mixed appearance spoke against its aesthetics or design features and compared it to “alien-like” appearance, for example, “that space thing on the head looks like something from outer space...”; “Looks like a space man or space woman...”; “Might help if he wasn't bald. Looks sort of alien without having any hair...” Interestingly, although many younger adults liked the human–machine blend in the mixed appearance, some older adults used the same reason to not like this appearance, for instance, “But this one, kind of gets the worst of both. It's not as pleasant as this one [robotic face] and not as familiar as this one [human face].” The

negative perceptions toward the mixed appearance could have arisen because of the aesthetic inconsistencies apparent in the superimposition of human–robotic features.

3.3.4 Preferences Within a Level of Human-Likeness—We could not conduct Chi-square tests to assess if younger and older adults differed in their most preferred human face, mixed face, and robotic face respectively because of the violation of an assumption in Chi-square analysis (i.e., the expected frequencies of each cell were not greater than five). However, informative trends and reasons for preferences were noted within each of the three levels of human-likeness.

Which Human Appearance was Preferred and Why?: All human faces were not evaluated similarly. A large majority of younger adults (84 %) and older adults (78 %) selected Female2 as the most preferred face for their robot in the human appearance category (Fig. 7). The most striking reason that participants gave for this preference was the robot gender. A majority of participants mentioned that they would prefer assistance from a female robot in their homes than a male robot. Some older adults also perceived this face to be like that of a nurse or a caregiver (e.g., “she looks like she’s a nurse. Got her hair back and prepared to do the work.”; “Whereas she looks like a caregiver, someone that could be in with you, close to you.”). On the other hand, some younger adults (both male and female) perceived a mother-like resemblance in the appearance (e.g., “She reminds me of my mom. So that’s kind of a deciding factor, it’s familiar.”; “especially if like the bathing thing, it seems like a motherly aspect, so it’d be more comfortable than like another guy, I guess, in the room.”).

The aesthetics of the face (e.g., the eyes, hairstyle) also influenced this selection. Although all faces had neutral expressions, some participants perceived a hint of smile on Female2’s face, which also affected their preference.

A few participants who, on the contrary, preferred a male face attributed more intelligence to a male-looking robot for investment-related tasks (decision-making). Other reasons for preferring a male face were perceptions of more strength in a male-robot, and in case of a few male participants, the comfort expected from a same-gender robot (e.g., a male older adult mentioned, “if they were going to bathe and be with me 24/7 I would want, probably a male. And if it wasn’t a robot, if it was a nurse, I’d probably pick a male nurse over a female nurse.”)

Which Robotic Appearance was Preferred and Why?: More than half of the younger adults preferred Nao over the other robot appearances (Fig. 8), seven of whom considered its appearance to be “cute”. Another primary reason given for the preference of Nao was its neutral expression which seemed to fit with how many younger adults imagine their robot to look like (e.g., “[Nao] just looks like a normal, when I picture a robot, that’s what I picture the white face with the eyes”; “[Nao] looks like very constan-t...like they don’t have as much emotion.”; “[Nao] seems like it’s like trying to be more robot-y...I like that [it] isn’t trying to be a human.”) Not having too many facial features was considered a plus, as can be observed in this comment, “[Nexi] is kind of reminiscent of a little brother or a little kid, and it has to do these chores, and I’d kind of feel bad. [Nao] doesn’t really have the ability to

move its mouth or raise its eyebrows, so I'd never know...which may not be good. Yeah. That feedback isn't there, which is kind of a plus."

Nao was not as popular among the older adults as it was among the younger adults. Forty-four percent of the older adults preferred Nexi over the other robot appearances. A common reason was the expressiveness of the appearance resulting from the combination of the facial features, as can be illustrated through this comment, "I like [Nexi's] big eyes. Uhh, I like his round face instead of that square face. I like the fact that he has a real mouth instead of a line. This looks like a brain center that he could be thinking about. His eyebrows look like they could move and make an expression on his face."

Which Mixed Appearance was Preferred and Why?: Younger and older adults differed in their most preferred mixed appearance (Fig. 9). Nexi + Female1 was most popular younger adults, but older adults showed higher preference for Pearl + Female2. Younger adults preferred Nexi + Female1 primarily because it was more human-looking than the other mixed-appearances. The younger age group also ascribed more intelligence to this appearance, which further led to their preference of this face.

In the older age group, 50 % participants preferred Pearl + Female2 and 37 % preferred Nexi + Female1. Pearl + Female2 was preferred primarily for the aesthetics and perceived personality (e.g., pleasant, companion-like) of the appearance. About half of the older adults who preferred Pearl + Female2 made specific comments about its eyes such as, "The eyes are telling me this robot can be trusted...there's just something about the eyes that just make, I mean when I communicate I look at a person's eyes...".

3.4 Do Perceptions of Robots of Different Levels of Human-Likeness Vary Across Tasks?

Participants' mean likeability ratings were generally between 2 (a little) and 3 (a fair amount) although younger adults' likeability for robot appearance exceeded 3 for chores and social task (Fig. 10). Younger adults seemed to like robot appearance more than human and mixed appearances for all tasks except decision-making. Older adults' likeability ratings for human and robot appearances seemed comparable for all tasks except decision-making for which likeability for robot appearance dropped considerably.

The univariate analysis of human-likeness X task on like-ability yielded a significant two-way interaction ($F(4.74, 293.91) = 8.27, p < 0.001, \eta_p^2 = 0.12$). The interaction was investigated further via post-hoc comparisons. When conditioned on task, and compared across human-likeness, paired *t* tests revealed that with the exception of the decision-making task, robotic appearance was liked more than the mixed appearance ($p < 0.001$). The human appearance was liked more than the mixed appearance for decision-making and social tasks ($p < 0.0167$). Additionally, robot appearance was liked more than the human appearance for assistance with a personal care task ($p = 0.01$).

Thus, one clear pattern observed across all DVs was that for the decision-making task, robotic appearance was not evaluated as positively and preferentially over mixed appearance as for other tasks. The qualitative data from the interview provided insights into why this was the case. People tended to use different evaluation criteria when judging appearance

preferences for a particular task. Specifically for decisionmaking, perceptions of intelligence and smartness in the appearance took precedence over perceptions of cuteness and friendliness. Example comments from participants (young adults) who preferred mixed appearance for decision making were: “it's just the whole intelligent look...”, “because he seems... wiser with the glasses”, and “that one looks more smart”. In the interview, almost half of the younger adults selected a mixed appearance as their most preferred appearance for decision-making. The proportion of older adults preferring a mixed appearance also increased for decision-making in comparison to other tasks, yet most older adults still tended to side with human or robotic appearances.

Similarly, participants focused on the “sociability” attribute when deciding their most preferred appearance for a social task. Thus, when asked to think specifically about a social task 60 % of older adults and 50 % of younger adults preferred human face for their robot. An older adult explained the reason for this preference, “because she's ah, well it's, she looks more capable of being sociable than these two. She looks more human-like.” A younger adult, who also preferred a human appearance for social task, commented, “particularly with a social task, you wanna be dealing with a human. Or at least, make it seem like you're dealing with a human more than a robot.”

Less distinct preferences for appearance emerged for assistance with chores. Because this task is both less critical and less interactive, appearance might have been deemed a less critical variable when considering robot assistance in this context. Robots, averaged across all appearances, were evaluated most positively (i.e., rated higher on likeability, trust, and PU) for assistance with chores, particularly in comparison with personal care and decision-making tasks.

For personal care, mean likeability rating for robotic appearance was higher than mixed and human appearance. However, this was not true for trust and PU ratings. During the interview, participants were divided in their preferences. Those who preferred a human appearance focused on the “care” aspect of the personal care task. Thus they associated more human-like care and capabilities with a human-looking robot: “I'm just more comfortable with this robot that looks more like a nurse or a nursing assistant. It looks like a humanoid that you could trust and I'm giving them the benefit of knowing how to aid and hold you as immerse into the water bathing or that sort of stuff”. On the contrary, many others were concerned about the “personal” aspect of the task. This group of participants did not want a human-looking robot to assist them with a task so private in nature, as is reflected in this comment, “sometimes personal care can get pretty involved, and I'd much rather have an impersonal looking creature caring for my personal needs.”

3.5 Do Young and Older Adults Differ in their Perceptions of Robots?

In the rating task, the three-way interaction of age \times human-likeness \times task was not significant for likeability ratings ($F(4.74, 293.91) = 0.73, p = 0.59, \eta_p^2 = 0.01$). The two-way interactions of age \times human-likeness ($F(1.48, 91.97) = 2.81, p = 0.08, \eta_p^2 = 0.04$) and age \times task ($F(2.45, 151.69) = 2.47, p = 0.08, \eta_p^2 = 0.04$) were also non-significant, probably because of low statistical power. Moreover, there was no main effect of age ($F(1, 62) = 0.12, p =$

0.73, $\eta_p^2=0.002$). However, age-cohort differences were found in the preference data gathered during the interview. These differences have been reported in previous sections.

In general, older adults tended to not prefer the mixed appearance whereas 25 % of the younger adults in the sample selected a mixed appearance as their most preferred appearance. Younger adults became even more receptive of the mixed appearance for the decision-making task whereas many older adults held on to their choice of a human or robotic appearance. Overall, more older adults favored a human appearance, primarily due to reasons of familiarity with the appearance, whereas younger adults tended to have a global preference for the more mechanical (i.e., more robotic) appearances. However, the heterogeneity in preferences within the age-groups was striking and should not be ignored.

4 Discussion

In human–human interaction, facial appearances influence formation of initial impressions, which in turn influence social behaviors. Similarly, in human–robot interaction, people's initial impressions of the robot may be based on its appearance, which may in turn affect their willingness to use the robot for the purpose it is designed for. One of the primary goals of this study was to investigate if initial perceptions formed towards robots would be influenced by the human-likeness of the robot's face, particularly when the robot is providing assistance with tasks that are traditionally carried out by humans. Moreover, although robots have the potential to help both younger and older adults, there is limited knowledge on how the two age groups' perceptions of robot human-likeness compare with each other. Therefore, an additional goal was to examine if younger and older adults differed in their perceptions.

At a general level, a mixed human–robot facial appearance was evaluated less positively than a highly human-looking or a highly robot-looking appearance. This trend was observed in the rating task across the measures of likeability, trust, and perceived usefulness for both younger and older adults. This finding seems aligned with the uncanny valley theory [17], implying that a robot face that partially imitates a human appearance evokes less positive perceptions than a more mechanical or a completely human-like robot face. However, one of the caveats of the earlier research on uncanny valley theory was the ill-defined context in which robot appearances were evaluated (e.g., [15,25]). This caveat was addressed in the current study by asking participants to imagine interacting with the robot in specific task contexts.

When the task was taken into account, the trends in perceptions were more complex and deviations from the uncanny valley pattern were observed. For example, robot (mechanical) appearance was evaluated more positively than the mixed appearance for chores, social, and personal care tasks. However, for decision-making task, mean ratings for robot appearance were comparable to those for the mixed appearance. Additionally, age-related differences in perceptions were noted. The younger adults (but not the older adults) evaluated the mixed appearance more positively for assistance with decision-making than with personal-care and social tasks. Thus, this study evaluated perceptions of a broader range of users and found differential perceptions across age.

Prior research on robot appearance that considered the robot's task did not assess the underlying reasons for the preference of one appearance over another (e.g., [36]). The multi-method approach used in the current study identified not only the patterns of perceptions across different appearances but also the possible reasons that influence the formation of such perceptions. For example, in the rating task participants' perceptions for the robot (mechanical) appearance were found to be least favorable for the decision-making task. The interview data revealed that participants varied their evaluation criteria for robot appearance across different tasks.

For the decision-making task, the appearance that evoked perceptions of intelligence, smartness, or wisdom was preferred for assistance. Perceptions of “cuteness” or “friendliness”, which were frequently mentioned as reasons for a general preference of the mechanical appearance, were not considered important when evaluating assistance for a cognitively demanding task such as decision-making. For a considerable proportion of the younger adults, a mixed appearance with an appropriate blend of human–mechanical appearance met the criterion of intelligence, and was preferred over the other appearances for decision-making. However, older adults considered the mixed appearances to be less familiar or alien-like, and were therefore, more in favor of the human appearance for this task.

The results of this study have implications both for advancing theoretical understanding of robot perceptions and for creating and applying guidelines for the design of robots. These are discussed separately in the next sections.

4.1 Theoretical Implications

As measured via the robot opinions questionnaire, both younger and older adults had generally positive opinions about using robots. However, compared to the younger adults, the older adults had less familiarity and experience with robots (assessed through the robot familiarity and use questionnaire). Therefore, the older adults' perceptions about a robot's appearance were more likely to be shaped by their expectations than by past experiences with a robot. Older adults' higher preference for a human appearance could be an outcome of such inexperience.

A primary reason for why human-looking robots might be favored over mechanical appearance is *familiarity* with the human appearance, particularly for performing tasks in the home that typically have been performed by humans [14]. However, such participants might assume the robot to be a perfect copy of a human, triggering the same nuance of familiarity as is evoked by another human. From the perspective of the uncanny valley theory [17] this would happen when the second peak of the graph is reached (see Fig. 1). The primary proposition of the uncanny valley theory is also based on the notion of familiarity. The argument is that as a robot is designed to appear more human-like, the familiarity with it reduces because the appearance seems to be a faulty replica of a human. However, if the human-like appearance can be perfected to reach a point where it is indistinguishable from human appearance, positive perceptions (due to high familiarity) will be evoked.

Although a majority of the older adults preferred a human appearance, another considerable proportion of them (37 %) leaned toward the least human-looking appearances for their robot. Many participants, including younger and older adults, raised concerns about not being able to differentiate a robot from a human. The closer the robot's face resembles to a human, the more likely it is to be anthropomorphized through overgeneralization effects [1,61]. Thus, it would be difficult for the human users to inhibit attributions of human strengths and weaknesses onto a human-looking robot. Moreover, people would already have expectations about the behavior of the robot. This would apply even to people who prefer a human appearance for their robot. For example, some participants assumed that a human-looking robot would be more capable than other robots, and therefore favored the appearance. On the contrary, some participants attributed human flaws (such as disobedience and betrayal) onto human-looking robots and therefore were more inclined toward a mechanical appearance. Therefore, familiarity-based overgeneralizations could be both beneficial and problematic for the acceptance of human-looking robots. Thus, theoretical models of robot perception need to be inclusive of the positive and negative effects of familiarity and expectations that emerge from a human appearance.

Robot appearance research has focused on identifying general patterns in perceptions and preferences. These patterns predict a trend of behavior for most people but ignore or undervalue inter-individual differences. For instance, in the current study, more positive evaluation of the human and the robot appearance over the mixed appearance offered support for the uncanny valley phenomenon at a nomothetic level. However, 25 % of the younger adult sample chose a mixed appearance as their *most preferred* appearance during the interview. Such participants might have a different trend of perceptions across varying levels of robot human-likeness. This means that the uncanny theory, even if validated at the group level, might not hold true at individual level.

One of the sources of the inter-individual differences can be the participants' age cohort. The present-day older and younger age groups differ not only in their direct experiences with robots but also in their exposures to robot-specific scientific fiction (e.g., novels, movies, and TV series). Such differences in experience could lead to different expectations toward robot appearance [39]. Thus, individual differences should also be incorporated into a model of robot perceptions by systematically considering a wider range of potential users.

Previous research on robot appearance may have underestimated *intra-individual* differences in appearance perceptions by not taking the context of human–robot interaction into account (e.g., [44]). The present study found that task context affected individuals' perceptions for robot human-likeness. Thus, individuals calibrated their appearance preferences based on the attributes of the task. Therefore, perceptions and preferences gauged in isolation with task contexts are less informative of the participants' evaluation criteria and judgment processes.

In addition, in the extant literature on humanoid appearance much emphasis has been on comparing the effects of different levels of human-likeness (e.g., [15,25]). The assumptions of such research undermine the variability *within* a particular level of human-likeness. It is not merely the degree of human-likeness but also specific characteristics such as gender,

expressiveness, aesthetics, and perceived capability or intelligence that influence people's perceptions. Moreover, tasks are also stereotyped by gender, which can further influence perceptions of male versus female looking robots [62]. People in the present study were more likely to prefer a female human-looking robot for general assistance in their homes. Assistance with chores, personal care, and social companionship are tasks stereotypically associated with females. However, participants' comments indicated that for more cognitively demanding tasks (e.g., decision-making in managing finances, other male gender-typed tasks), preference for assistance might shift toward a male-looking robot. Additionally, some male participants leaned toward a male-robot for assistance with personal care tasks.

Robotic (mechanical) appearance was preferred overall when it was perceived as “cute”, “friendly”, “trustworthy” and/or easy to command. However, all mechanical appearances were not perceived equally favorably. Similarly, even though the mixed appearance was less positively evaluated, there were differences in perceptions within that category. Therefore, specific characteristics of any robot appearance (aesthetics/features, expressiveness, gender, etc.) may interact with the robot's human-likeness to affect robot acceptance.

Based on the response on robot facial appearance questionnaire, which asked for opinions about robot appearance in the context of home use (but not for any specific task), we found that both the age-groups considered eyes and mouth highly desirable features in a robot's face whereas nose and ears were desired less strongly. Overall, older adults seemed to want more features than did younger adults (e.g., hair on the head, lips, soft skin). In fact, younger adults *did not* want their robot to have hair on its head and their responses were significantly different from the older adults in the opposite direction. Consistent with this is the observation that Nexi + Female1 (see Fig. 2), which is noticeably bald, was the most preferred mixed appearance for majority of younger adults (Fig. 9).

Older adults were more likely to want their robot to be completely human-looking whereas younger adults tended to be neutral in their response to this question, which is possibly why older adults also desired more facial features in the robot's face. Moreover, although the older adults were neutral about the presence of male or features in the robot's face, the younger cohort showed significant inclination for female features. However, both the age groups wanted their robot to have a face that is round (older adults desired it even more than younger adults), unique, attractive, and expressive.

4.2 Applied Implications

From the applied perspective, it is important to consider both younger and older adults as heterogeneous groups. On average, older adults may show a higher inclination toward human-like appearance of robots; however, differences in preferences *within* the two age groups should not be overlooked. A considerable proportion of older adults also preferred robotic (or mechanical appearance). Younger adults' preferences were even more diverse—half of them preferred robotic appearance, a fourth preferred human appearance and the remaining fourth preferred mixed appearance. Thus, in general, there is wide variability in people's attitudes toward mechanical and humanoid appearances. Overall, robot designers

should consider designing robot appearances keeping the following four categories of users in mind (Fig. 11):

- *Category 1* Those who prefer a mechanical appearance
- *Category 2* Those who prefer some humanoid characteristics (i.e., a mixed appearance).
- *Category 3* Those who prefer a human appearance
- *Category 4* Those with no strong preference for appearance

The primary reasons associated with a preference for human-like appearance (category 3) are familiarity, ease of interaction, perceptions of higher capability, and expectations of companionship from the robot. As was noted in the present study, people in this category also consider human-looking robots as the most “developed” kind of robots and expect the robot to be more than a mechanical tool.

In contrast, individuals who do not want their robot to be human-like (categories 2 and 3) may be concerned about confusing a robot with a human being. Some of them even tend to attribute human flaws (such as disobedience and betrayal) to human-looking robots. Category 2 comprises individuals who do not want the robot to be highly human-looking but would not mind some humanoid features particularly if the robot appears “cute”, “friendly”, “trustworthy” and/or “easy to command”. Thus, besides the degree of human-likeness, another important consideration in robot face design is about the composites of the face that bring about an expression or an apparent character. People in categories 1 and 4 view a robot only as a mechanical tool. However, what distinguishes the two categories is that category 1 people want their robot to look like a machine whereas category 4 are not concerned about the looks of the robot.

When asked to think about using robots in specific task scenarios, people's evaluation criteria for an appearance (e.g., cute vs. smart) varied across tasks. Thus, if the robot is designed to specifically assist the user with critical decisions, the robot should not be given a funny demeanor, but an intelligent look. For assistance with a personal care task (e.g., bathing), users can be divided into two broad groups: those who would prefer the human appearance and those who would prefer a mechanical/robotic appearance. Human appearance is preferred for assistance with personal care because it evokes in the user perceptions of “nurse-like” capabilities. Thus, people may be more comfortable in taking assistance from the robot. However, the human appearance might also be considered invasive for a highly personal task, which is why a mechanical appearance would be preferred by the other user group. For a social task (chatting, providing companionship, helping learn a new skill, etc.) human-like sociability would be an important attribute of the robot appearance.

A robot's appearance sets expectations in the user about its capabilities. Therefore, for the robot to be accepted, it is important that its actual capabilities do not violate the set expectations. People who want their robot to look like a human also have high expectations from the robot in terms of its abilities to behave like a human. However, if the robot is not able to keep up with such high expectations, it may be more strongly rejected in comparison

to a less human-like or non-human-like robot that evoked lower expectations about its capabilities [63].

4.3 Methodological and Measurement Considerations

This study was designed to systematically investigate the effect of robot human-likeness on peoples' perceptions. One strength was the use of multiple pictures for every level of human-likeness as it provided multiple data points at each level of robot human-likeness. Thus, we could assess if participants were uniform in their perceptions within every level of appearance, and we found that they were not. For instance, participants had highly positive perceptions toward one of the human faces but not toward the others. This supports that it is not only the human-likeness of the face but also the specific characteristics of the human appearance that impact perceptions.

Another strong aspect of the study was the assessment of robot appearance perceptions in the context of tasks. Robots can be designed to assist with various activities of daily living; however depending on the activity, the nature of robots' assistance would differ in terms of physical interaction with the user (e.g., proximal interactions for personal care vs. distal for chores), cognitive support needed (high for decision-making vs. low for chores), and social interactions (high for social task vs. low for chores). Thus, people's perceptions of robot human-likeness would also likely be dependent on the nature of the assistance needed from the robot.

The use of multiple dependent measures was also an important extension of earlier HRI studies. Evaluating perceptions across different measures provided a clearer relation among the constructs that constitute perceptions and predict attitudinal acceptance of robots. Although the three measures of positive evaluations were highly positively correlated, perceived usefulness and trust were almost perfectly correlated for older adults. This implies that the robots that are considered more useful by older adults may also be trusted more and vice-versa. An alternative possibility is that people do not differentiate those terms and thus, these measures are not sensitive enough to differentiate attitudes on those dimensions.

Despite the systematic design, the study was not without limitations. The biggest caveat was the lack of robot interactivity due to the use of still pictures. Moreover, although vignettes were provided for every task situation, participants were restricted to their imaginations to evaluate the nature of robot assistance with different tasks. In addition, the results of the rating task did not allow us to disambiguate effects of human-likeness from other factors such as facial features, perceived gender, and perceived expressions on the robot's face. However, the interview provided insights into the role of these variables in perception formation.

The use of morphed pictures to create an in-between level of human-humanoid appearance was also a challenge. Although mixed appearances were represented by four morphed pictures, the morphed faces were probably not of the same aesthetic quality. Thus, the perception of morphed images might also be based on the aesthetical consistency in the merged features whereby some morphs maybe more pleasant-looking than others [64].

Another limitation was the use of single items as measures of likeability, trust, perceived usefulness, and anxiety. Multi-item measures could have been more informative but were not practical based on considerations for a reasonable length of the rating task. However, problems were noted in using “anxiety” as a dependent measure of negative perception. As was expected, anxiety was negatively correlated with perceived usefulness, trust, and likeability (see Table 3). However, less confidence can be put into the strength of the correlations due to the issues associated with the anxiety data. Future studies should use such constructs to measure negative attitudes toward robots that are more easily and unambiguously understood by participants (e.g., fear, discomfort). Moreover, using bipolar Likert-type items, with a neutral-point in the middle, would allow for the measurement of positive to negative valence on the same dimension (e.g., comfort–discomfort, like–dislike, trust–distrust, useful–useless). In addition, it would minimize confusions regarding the direction of the scale.

The findings of this study should be carefully generalized. The older age group sample in the study was represented by relatively healthy and educated older adults living in the Atlanta metropolitan area of the United States. Older adults with different backgrounds from the present sample might have different preferences and attitudes toward robots. Similarly, the younger age group was represented by current undergraduate students. The interactive effect of educational background and age-cohort experiences on attitudes toward robots was not explored in the current study. Therefore, the degree to which less educated younger adults would vary in their perceptions of and preferences toward robot appearance remains unknown. Moreover, future studies should focus on a broader range of users including middle-aged adults.

5 Conclusion and Future Directions

People's perceptions of robot faces clearly varied as a function of robot human-likeness. In general, people perceived a mixed human–robot appearance less favorably compared to highly human and more robotic appearance. Additionally the nature of task also influenced people's overall perceptions of robots. Robots were most positively evaluated for assistance with chores and less positively for personal care and decision-making. This finding is consistent with previous studies (e.g., [34]). Moreover, task and robot human-likeness had an interactive effect on people's likeability, trust, and perceived usefulness toward robots.

There were age-related differences in *preferences* of robot human-likeness. Older adults showed a higher inclination toward human-looking appearance of robots whereas younger adults' preferences were more distributed across the levels of human-likeness. An appearance with mixed human–robot features may be more likely to be rejected by older adults than by younger adults, and the difference would be most striking for a decision-making task.

Besides the human-likeness of the robot face, perceptions of robot appearances were also influenced by other factors such as robot gender, facial features/aesthetics, expressiveness, perceived personality, and perceived capability. Future studies should measure the relative weight of these different factors in the formation of perceptions, both at a global level and at

a task-specific level. Moreover, an understanding of the aesthetics that evoke perceptions of intelligence, sociability, and/or human-likeness could further improve the design of robots.

Acknowledgments

This study was part of the first author's master's thesis [59]. The research was supported in part by a grant from the National Institutes of Health (National Institute on Aging) Grant P01 AG17211 under the auspices of the Center for Research and Education on Aging and Technology Enhancement (CREATE; <http://www.create-center.org>). We would like to thank Dr. Leslie DeChurch and Dr. Arthur D. Fisk for their input at various stages of this study. We also acknowledge the assistance from Laura Matalenas, Russell Bowles, and Shelby Long with data collection and analysis. Portions of data from this research study have been presented previously in Prakash and Rogers [58].

References

1. Zebrowitz LA, Montepare JM. Social psychological face perception: why appearance matters. *Soc Pers Psychol Compass*. 2008; 2(3):1497–1517.
2. Zebrowitz, LA. *Reading faces: window to the soul?*. Westview; Boulder, CO: 1997.
3. Masip J, Garrido E, Herrero C. Facial appearance and impressions of ‘credibility’: the effects of facial babyishness and age on person perception. *Int J Psychol*. 2004; 39(4):276–289.
4. Bar M, Neta M, Linz H. Very first impressions. *Emotion*. 2006; 6(2):269. [PubMed: 16768559]
5. Zebrowitz LA, Rhodes G. Sensitivity to “bad genes” and the anomalous face overgeneralization effect: cue validity, cue utilization, and accuracy in judging intelligence and health. *J Nonverbal Behav*. 2004; 28(3):167–185.
6. Hassin R, Trope Y. Facing faces: studies on the cognitive aspects of physiognomy. *J Pers Soc Psychol*. 2000; 78(5):837–852. [PubMed: 10821193]
7. Brownlow S. Seeing is believing: facial appearance, credibility, and attitude change. *J Nonverbal Behav*. 1992; 16(2):101–115.
8. Riggio RE, Widaman KF, Tucker JS, Salinas C. Beauty is more than skin deep: components of attractiveness. *Basic Appl Soc Psych*. 1991; 12(4):423–439.10.1207/s15324834basps1204_4
9. Walster E, Aronson V, Abrahams D, Rottman L. Importance of physical attractiveness in dating behavior. *J Pers Soc Psychol*. 1966; 4(5):508. [PubMed: 6008393]
10. Todorov A, Mandisodza AN, Goren A, Hall CC. Inferences of competence from faces predict election outcomes. *Science*. 2005; 308(5728):1623–1626. [PubMed: 15947187]
11. Reeves, B.; Nass, C. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press; New York: 1996.
12. Nass C, Moon Y. Machines and mindlessness: social responses to computers. *J Soc Issues*. 2000; 56(1):81–103.
13. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q*. 1989; 13(3):319–340.
14. Blow, M.; Dautenhahn, K.; Appleby, A.; Nehaniv, CL.; Lee, D. 1st ACM SIGCHI/SIGART conference on human–robot interaction, 2006. ACM; 2006. The art of designing robot faces: dimensions for human-robot interaction; p. 331-332.
15. MacDorman KF, Ishiguro H. The uncanny advantage of using androids in cognitive and social science research. *Interact Stud*. 2006; 7(3):297–337.
16. DiSalvo, CF.; Gemperle, F.; Forlizzi, J.; Kiesler, S. 4th Conference on designing interactive systems: processes, practices, methods, and techniques, 2002. ACM; 2002. All robots are not created equal: the design and perception of humanoid robot heads; p. 321-326.
17. Mori M. The uncanny valley. *Energy*. 1970; 7(4):33–35. translated into English by MacDorman KF, Minato T, 2005, Proceedings of the humanoids-2005 workshop: views of the uncanny valley, Tsukuba, Japan.
18. Mori M, MacDorman KF, Kageki N. The uncanny valley [from the field]. *Robot Automat Mag*. 2012; 19(2):98–100.
19. Chen H, Russell R, Nakayama K, Livingstone M. Crossing the “uncanny valley”: adaptation to cartoon faces can influence perception of human faces. *Percept*. 2010; 39(3):378.

20. Geller T. Overcoming the uncanny valley. *IEEE Comput Graphics Appl.* 2008; 28(4):11–17.
21. Groom V, Nass C, Chen T, Nielsen A, Scarborough JK, Robles E. Evaluating the effects of behavioral realism in embodied agents. *Int J Hum–Comput Stud.* 2009; 67(10):842–849.
22. MacDorman KF, Green RD, Ho CC, Koch CT. Too real for comfort? Uncanny responses to computer generated faces. *Comput Hum Behav.* 2009; 25(3):695–710.
23. Schneider E, Wang Y, Yang S. Exploring the uncanny valley with Japanese video game characters. *Proceedings of DiGRA conference 2007.* 2007:546–549.
24. Thompson JC, Trafton JG, McKnight P. The perception of humanness from the movements of synthetic agents. *Percept.* 2011; 40(6):695–704.
25. MacDorman KF. Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: an exploration of the uncanny valley. *ICCS/CogSci-2006 long symposium: toward social mechanisms of android science, 2006.* 2006:26–29.
26. Rosenthal-von der Pütten AM, Krämer NC. How design characteristics of robots determine evaluation and uncanny valley related responses. *Comput Hum Behav.* 2014; 36:422–439.
27. Burleigh TJ, Schoenherr JR, Lacroix GL. Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Comput Hum Behav.* 2013; 29(3):759–771.
28. Smarr, CA.; Fausset, CB.; Rogers, WA. Understanding the potential for robot assistance for older adults in the home environment (HFA-TR-1102). Georgia Institute of Technology, School of Psychology, Human Factors and Aging Laboratory; Atlanta, GA: 2011.
29. Lawton MP. Aging and performance of home tasks. *Hum Factors.* 1990; 32(5):527–536. [PubMed: 2074107]
30. Rogers WA, Meyer B, Walker N, Fisk AD. Functional limitations to daily living tasks in the aged: a focus group analysis. *Hum Factors.* 1998; 40(1):111–125.10.1518/001872098779480613 [PubMed: 9579107]
31. Fisk, AD.; Rogers, WA.; Charness, N.; Czaja, SJ.; Sharit, J. Designing for older adults: principles and creative human factors approaches. 2nd. CRC; Boca Raton, FL: 2009.
32. Broadbent E, Tamagawa R, Patience A, Knock B, Kerse N, Day K, MacDonald BA. Attitudes towards health-care robots in a retirement village. *Australas J Ageing.* 2012; 31(2):115–120.10.1111/j.1741-6612.2011.00551.x [PubMed: 22676171]
33. Prakash, A.; Beer, JM.; Deyle, T.; Smarr, CA.; Chen, TL.; Mitzner, TL.; Kemp, CC.; Rogers, WA. 8th ACM/IEEE international conference on human–robot interaction (HRI). IEEE; 2013. Older adults' medication management in the home: how can robots help?; p. 283-290.
34. Smarr CA, Mitzner TL, Beer JM, Prakash A, Chen TL, Kemp CC, Rogers WA. Domestic robots for older adults: attitudes, preferences, and potential. *Int J of Soc Robot.* 2014; 6(2):229–247. [PubMed: 25152779]
35. Lohse M, Hegel F, Swadzba A, Rohlfing K, Wachsmuth S, Wrede B. What can I do for you? Appearance and application of robots. *Workshop on The Reign of Catz and Dogz? The role of virtual creatures in a computerised society.* 2007:121–126.
36. Goetz J, Kiesler S, Powers A. Matching robot appearance and behavior to tasks to improve human-robot cooperation. *Proceedings of the 12th IEEE international workshop on robot and human interactive communication (RO-MAN), 2003.* 2003:55–60.10.1109/ROMAN.2003.1251796
37. Hinds PJ, Roberts TL, Jones H. Whose job is it anyway? A study of human–robot interaction in a collaborative task. *Hum–Comput Interact.* 2004; 19(1):151–181.
38. Beer, JM.; Smarr, C.; Chen, TL.; Prakash, A.; Mitzner, TL.; Kemp, CC.; Rogers, WA. 7th ACM/IEEE international conference on human–robot interaction (HRI), 2012. IEEE; 2012. The domesticated robot: design guidelines for assisting older adults to age in place; p. 335-342.
39. Ezer, N. Dissertation. Georgia Institute of Technology; 2008. Is a robot an appliance, teammate, or friend? Age-related differences in expectations of and attitudes towards personal home-based robots.
40. Wu YH, Fassett C, Rigaud AS. Designing robots for the elderly: appearance issue and beyond. *Arch Gerontol Geriatr.* 2012; 54(1):121–126. [PubMed: 21349593]

41. Walters ML, Syrdal DS, Dautenhahn K, te Boekhorst R, Koay KL. Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Auton Robot*. 2008; 24(2):159–178.
42. Nomura T, Kanda T, Suzuki T, Kato K. Psychology in human–robot communication: an attempt through investigation of negative attitudes and anxiety toward robots. 13th IEEE international workshop on robot and human interactive communication (RO-MAN), 2004. 2004:35–40.
43. Hegel, F.; Lohse, M.; Wrede, B. 18th IEEE international symposium on robot and human interactive communication (RO-MAN), 2009. IEEE; 2009. Effects of visual appearance on the attribution of applications in social robotics; p. 64–71.
44. Bartneck, C.; Kanda, T.; Ishiguro, H.; Hagita, N. RO-MAN, 2009. IEEE; 2009. My robotic doppelgänger—a critical look at the uncanny valley; p. 269–276.
45. Mathur, MB.; Reichling, DB. 4th ACM/IEEE international conference on human–robot interaction (HRI), 2009. ACM; 2009. An uncanny game of trust: social trustworthiness of robots inferred from subtle anthropomorphic facial cues; p. 313–314.
46. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag Sci*. 2000; 46(2):186–204.
47. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Quart*. 2003; 27(3):425–478.
48. Bagozzi RP, Davis FD, Warshaw PR. Development and test of a theory of technological learning and usage. *Human Relat*. 1992; 45(7):659–686.10.1177/001872679204500702
49. Goodhue DL, Thompson RL. Task-technology fit and individual performance. *MIS Quart*. 1995; 19(2):213–236.
50. Walters, ML. Dissertation. University of Hertfordshire; 2008. The design space for robot appearance and behaviour for social robot companions.
51. Nomura T, Suzuki T, Kanda T, Kato K. Measurement of negative attitudes toward robots. *Interact Stud*. 2006; 7(3):437–454.
52. Nomura, T.; Suzuki, T.; Kanda, T.; Kato, K. Proceedings of the 15th IEEE international symposium on robot and human interactive communication (RO-MAN), 2006. IEEE; 2006. Measurement of anxiety toward robots; p. 372–377.
53. Broadbent, E.; Tamagawa, R.; Kerse, N.; Knock, B.; Patience, A.; MacDonald, BA. Proceedings of RO-MAN, 2009. IEEE; 2009. Retirement home staff and residents' preferences for healthcare robots; p. 645–650.
54. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. *Int J of Soc Robot*. 2010; 2(4):361–375.10.1007/s12369-010-0068-5
55. Weiss, HM.; Cropanzano, R. Affective events theory: a theoretical discussion of the structure, causes and consequences of affective experiences at work. In: Staw, BM.; Cummings, LL., editors. *Research in organizational behavior: an annual series of analytical essays and critical reviews*. Vol. 18. Elsevier Science/JAI; New York: 1996. p. 1–74.
56. Lewis JD, Weigert A. Trust as a social reality. *Soc Forces*. 1985; 63(4):967–985.
57. McAllister DJ. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Acad Manage J*. 1995; 38(1):24–59.
58. Prakash A, Rogers WA. Younger and older adults' attitudes toward robot faces effects of task and humanoid appearance. *Human factors and ergonomics society 57th annual meet*, 2013. 2013:114–118.
59. Prakash, A. Master's Thesis. Georgia Institute of Technology; 2013. Understanding younger and older adults' perceptions of humanoid robots: effects of facial appearance and task.
60. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, Sharit J. Factors predicting the use of technology: findings from the Center for Research and Education on aging and technology enhancement (CREATE). *Psychol Aging*. 2006; 21(2):333–352. [PubMed: 16768579]
61. Epley N, Waytz A, Cacioppo JT. On seeing human: a three-factor theory of anthropomorphism. *Psychol Rev*. 2007; 114(4):864–886. [PubMed: 17907867]
62. Heilman ME, Wallen AS, Fuchs D, Tamkins MM. Penalties for success: reactions to women who succeed at male gender-typed tasks. *J App Psychol*. 2004; 89(3):416–427.

63. Paepcke, S.; Takayama, L. Proceedings of the 5th ACM/IEEE international conference on human-robot interaction (HRI), 2010. IEEE; 2010. Judging a bot by its cover: an experiment on expectation setting for personal robots; p. 45-52.
64. Hanson, D.; Olney, A.; Prilliman, S.; Mathews, E.; Zielke, M.; Hammons, D.; Fernandez, R.; Stephanou, H. Proceedings of the national conference on artificial intelligence, 2005, Menlo Park, CA. MIT Press; Cambridge, MA: 2005. Upending the uncanny valley; p. 1728-1729.

Abbreviations

| | |
|--------------|-----------------------------------------|
| ADLs | Activities of daily living |
| IADLs | Instrumental activities of daily living |
| EADLs | Enhanced activities of daily living |
| DVs | Dependent variables |
| PU | Perceived usefulness |

Biographies

Akanksha Prakash is an Engineering Psychology Ph.D. student at the Georgia Institute of Technology. She earned a Bachelor of Technology degree in Electrical Engineering from the Indian Institute of Technology (IIT), Roorkee, India, in 2007, and a M.S. in Engineering Psychology from Georgia Tech in 2013. Her research interests include human-robot interaction, social networks and connectedness, social cognition, and memory.

Wendy A. Rogers is Professor of Psychology and Coordinator of the Engineering Psychology Program at the Georgia Institute of Technology. Her research interests include cognitive aging, automation, skill acquisition, human-robot interaction, and technology acceptance. She is Director of the Human Factors and Aging Laboratory (<http://www.hfaging.org>), funded by the National Institutes of Health (National Institute on Aging) through the Center for Research and Education on Aging and Technology Enhancement (CREATE <http://www.create-center.org>) and the Department of Education through the Rehabilitation and Engineering Research Center on Technologies for Successful Aging with Disabilities (TechSAge <http://www.techsage.gatech.edu>).

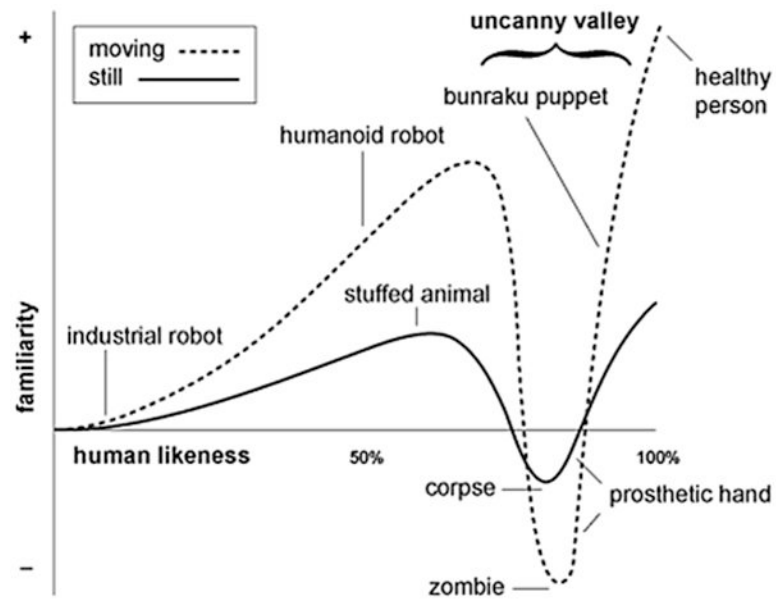


Fig. 1. Mori's hypothesized uncanny valley diagram (translated by MacDorman and Minato [17])

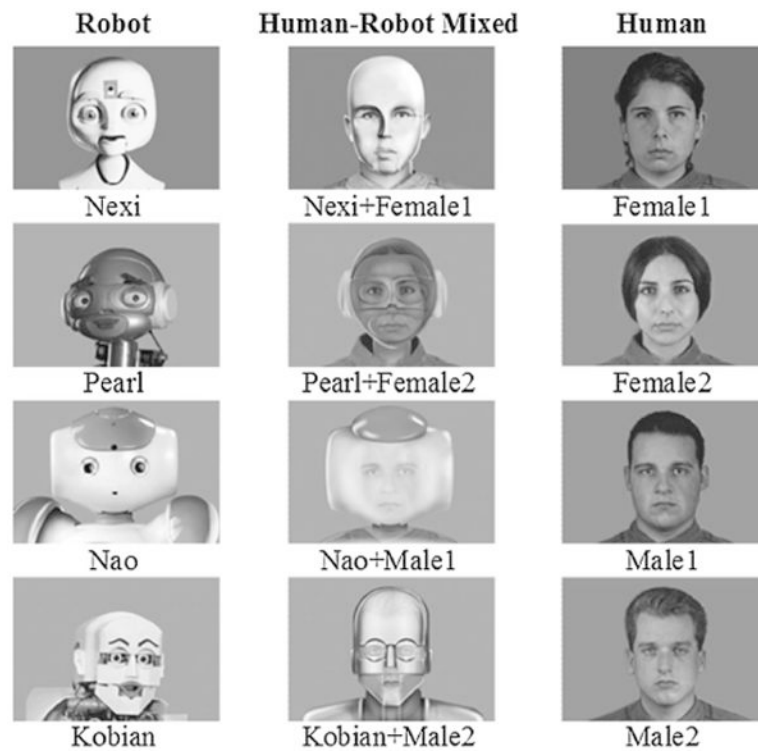


Fig. 2. Pictures used in the study to represent different levels of human-like appearance

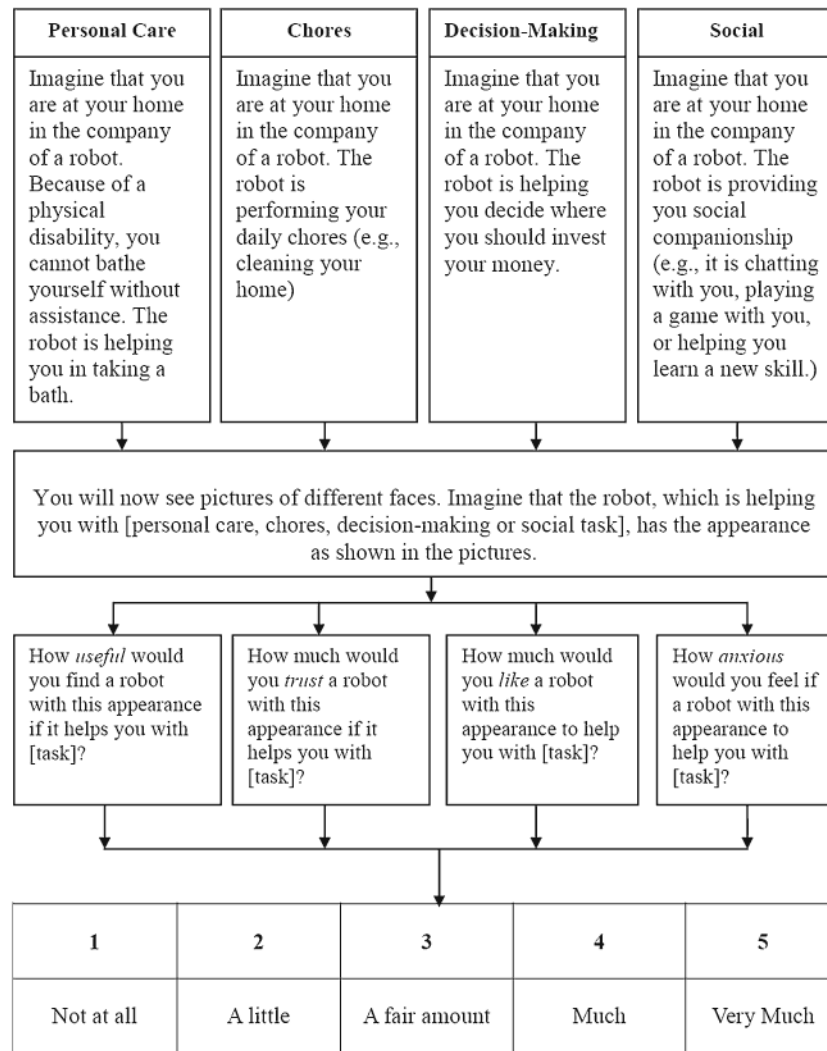


Fig. 3. Schematic representation of the rating tasks

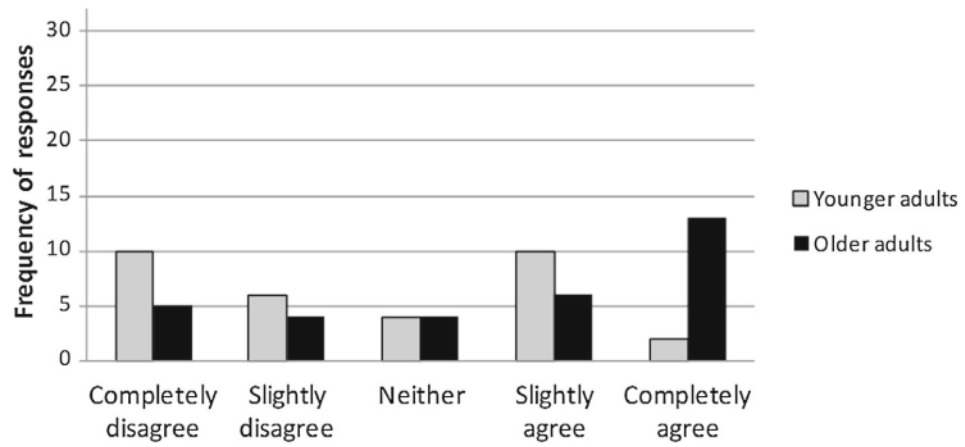


Fig. 4. Frequency distribution of responses on the item “I would want my robot to look exactly like a human”

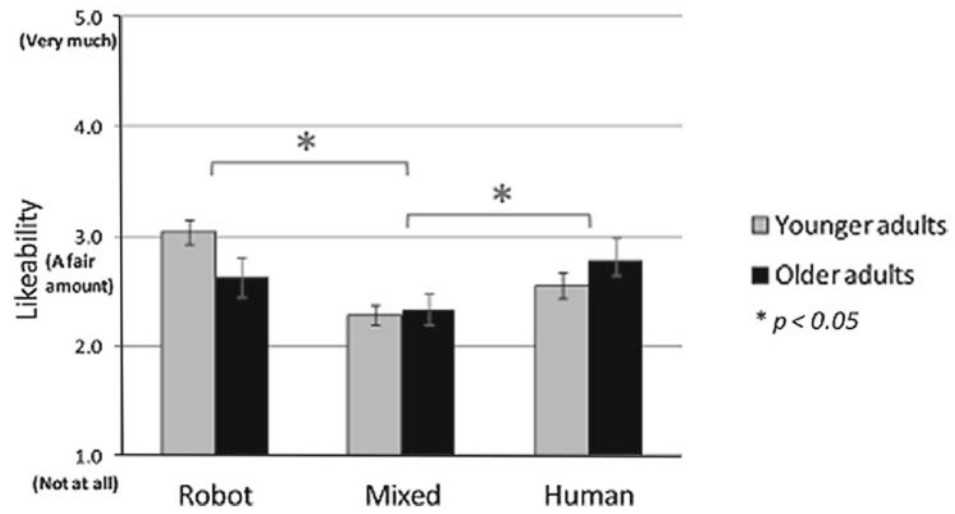


Fig. 5.
Mean likeability ratings by human-likeness. Error bars represent standard error of mean

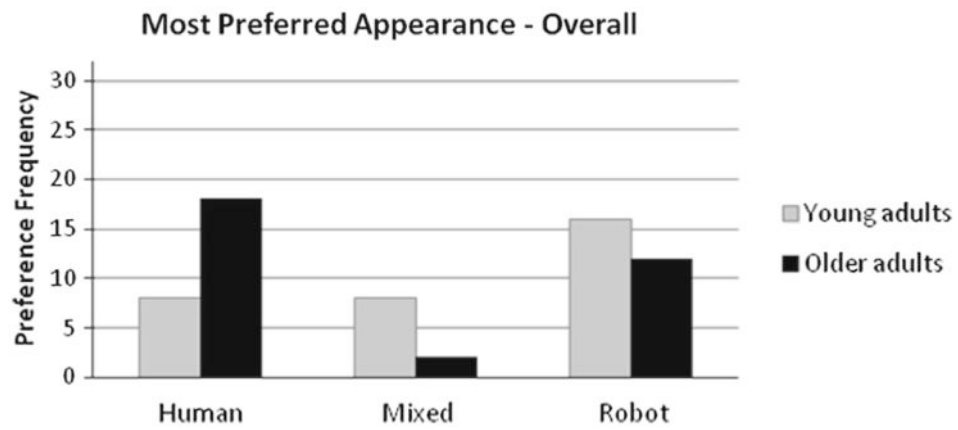


Fig. 6. Frequency distribution of younger and older adults' most preferred face for their robot

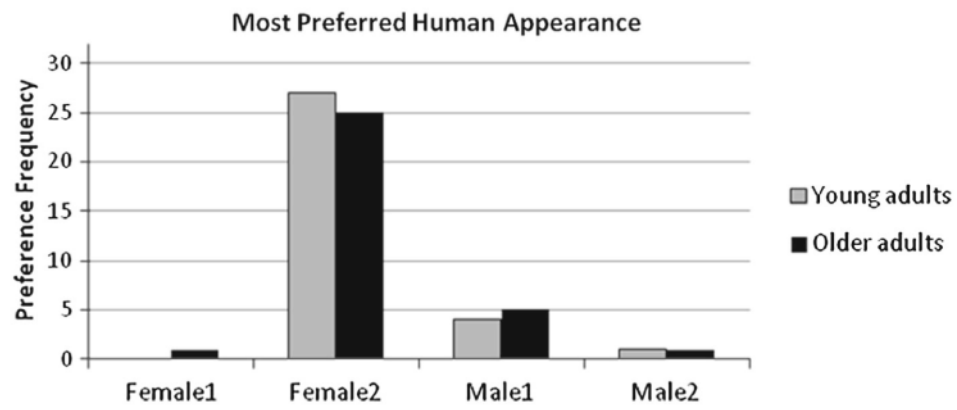


Fig. 7. Frequency distribution of younger and older adults' most preferred human appearance

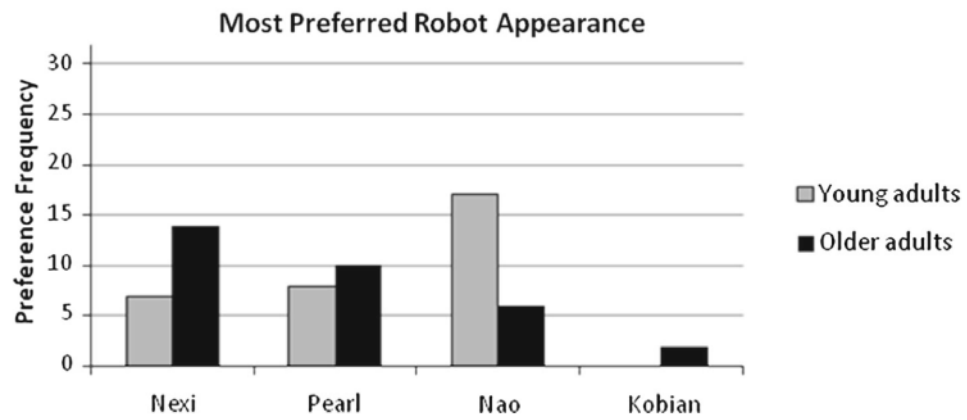


Fig. 8. Frequency distribution of younger and older adults' most preferred robot appearance

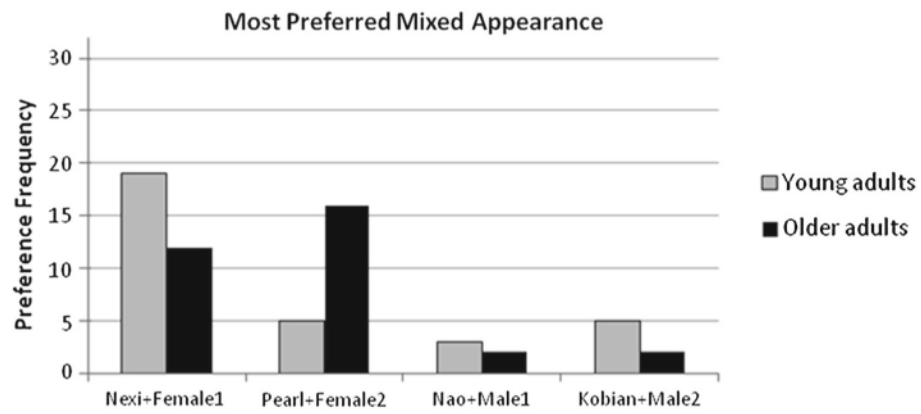


Fig. 9. Frequency distribution of younger and older adults' most preferred mixed appearance

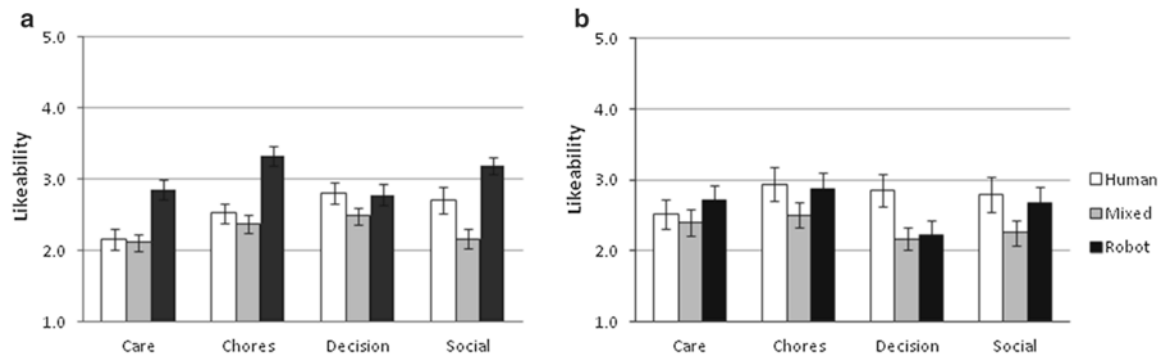


Fig. 10.

Mean likeability ratings by age-group, human-likeness, and task. Error bars represent standard error of the mean. Scale anchors were 1 (not at all), 3 (a fair amount), and 5 (very much). **a** Younger adults; **b** older adults

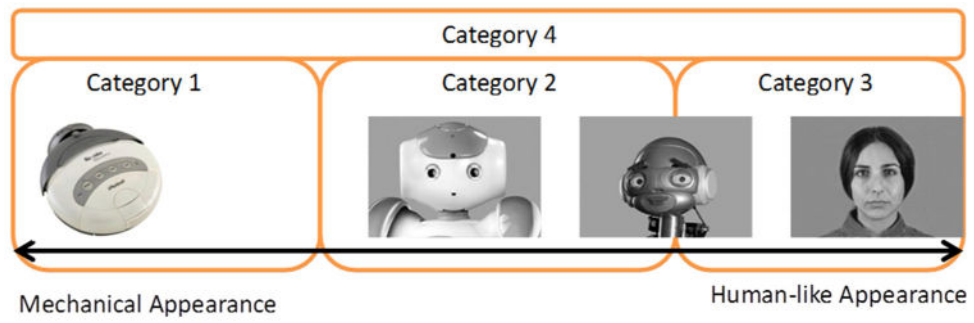


Fig. 11. Depiction of four categories of users based on preferences for robot appearances, from mechanical to human-like

Table 1
Robot opinions and robot familiarity and use questionnaires

| | Younger adults (N = 32) | | Older adults (N = 32) | | <i>t</i> test |
|----------------------------------------|-------------------------|-----------|-----------------------|-----------|------------------------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>t</i> (62) <i>p</i> |
| Robot opinions ^a | | | | | |
| Perceived usefulness | 5.37 | 0.97 | 5.30 | 1.38 | 0.24 0.81 |
| Perceived ease of use | 5.25 | 1.28 | 5.34 | 0.99 | -0.29 0.77 |
| Robot familiarity and use ^b | 1.77 | 0.23 | 1.54 | 0.41 | 2.73 0.01* |
| Autonomous car | 1.84 | 0.92 | 1.13 | 0.83 | 3.23 <0.01* |
| Domestic/home robot | 2.22 | 0.87 | 1.78 | 0.75 | 2.15 0.03* |
| Entertainment/toy robot | 3.03 | 0.82 | 1.78 | 0.91 | 5.78 <0.01* |
| Manufacturing robot | 2.19 | 0.40 | 1.88 | 0.55 | 2.60 0.01* |
| Military robot | 1.94 | 0.25 | 1.81 | 0.54 | 1.20 0.23 |
| Personal robot 2 (PR2) | 1.00 | 0.67 | 1.31 | 0.82 | -1.67 0.10 |
| Remote presence robot | 0.84 | 0.63 | 0.97 | 0.78 | -0.70 0.48 |
| Research robot | 1.69 | 0.64 | 1.41 | 0.71 | 1.66 0.10 |
| Robot lawn mower | 1.53 | 0.62 | 1.56 | 0.67 | -0.19 0.85 |
| Robot security guard | 1.28 | 0.63 | 1.19 | 0.65 | 0.76 0.45 |
| Space exploration robot | 2.00 | 0.00 | 1.81 | 0.47 | 2.25 0.03* |
| Surgical robot | 1.84 | 0.37 | 1.78 | 0.66 | 0.47 0.64 |
| Unmanned aerial vehicle | 1.63 | 0.66 | 1.69 | 0.69 | -0.37 0.71 |

^a Scale is 1 (=extremely unlikely) to 7 (=extremely likely) regarding questions of usefulness and ease of use

^b Scale is 0 (=not sure what it is) to 4 (=have used or operated it frequently). Mean is across the 13 different kinds of robots

* Significant at 0.05 level

Table 2

Facial appearance questionnaire

| Facial features/ Characteristics | Younger adults (N = 32) | | | Older adults (N = 32) | | | Wilcoxon rank– sum test of age- group difference | | |
|----------------------------------|--------------------------------------|---------|--------|--------------------------------------|---------|--------|--------------------------------------------------|--------|---|
| | One-sample Wilcoxon signed rank test | | Median | One-sample Wilcoxon signed rank test | | Median | One-sample Wilcoxon signed rank test | | p |
| | Z | p | | Z | p | | Z | p | |
| Face | 5.00 | <0.001* | 5.00 | -4.27 | <0.001* | 5.00 | -0.26 | 0.80 | |
| Eyes | 5.00 | <0.001* | 5.00 | -4.46 | <0.001* | 5.00 | -1.06 | 0.29 | |
| Nose | 3.00 | 0.01* | 4.00 | -3.40 | 0.001* | 4.00 | -1.73 | 0.08 | |
| Mouth | 5.00 | <0.001* | 5.00 | -4.45 | <0.001* | 5.00 | 0.00 | 1.00 | |
| Lips | 3.00 | 0.06 | 5.00 | -4.28 | <0.001 | 5.00 | -3.35 | 0.001* | |
| Ears | 3.00 | 0.02* | 4.00 | -3.37 | 0.001* | 4.00 | -1.79 | 0.07 | |
| Hair on head | 2.00 | 0.02* | 3.00 | -1.93 | 0.05 | 3.00 | -3.41 | 0.001* | |
| Soft skin | 3.00 | 0.77 | 3.00 | -2.48 | 0.01* | 3.00 | -2.03 | 0.04* | |
| Round face | 4.00 | <0.001* | 3.50 | -2.60 | 0.01* | 3.50 | -2.40 | 0.02* | |
| Exactly human-like | 2.50 | 0.09 | 4.00 | -1.97 | 0.048* | 4.00 | -2.61 | 0.01* | |
| Female features | 4.00 | 0.002* | 3.00 | -0.69 | 0.49 | 3.00 | -1.14 | 0.25 | |
| Male features | 3.00 | 0.22 | 3.00 | -1.17 | 0.24 | 3.00 | -0.31 | 0.75 | |
| Uniqueness | 4.00 | 0.02* | 4.00 | -3.35 | 0.001* | 4.00 | -0.73 | 0.46 | |
| Attractive-ness | 4.00 | <0.001* | 4.00 | -3.86 | <0.001* | 4.00 | -1.10 | 0.27 | |
| Expressive-ness | 4.00 | <0.001* | 4.00 | -3.78 | <0.001* | 4.00 | -0.61 | 0.54 | |

Items are listed in the order in which they appeared in the questionnaire. Scale is 1 (=completely disagree) to 5 (=completely agree).

* Significant at 0.05 level

Table 3

Correlations of dependent measures

| Age | Variable | M | SD | Likeability | Trust | Perceived usefulness |
|----------------|----------------------|------|------|-------------|---------|----------------------|
| Younger adults | Likeability | 2.63 | 0.36 | | | |
| | Trust | 2.80 | 0.37 | 0.68** | | |
| | Perceived usefulness | 2.99 | 0.42 | 0.66** | 0.88** | |
| | Anxiety | 2.67 | 0.58 | -0.33 | -0.52** | -0.39* |
| Older adults | Likeability | 2.58 | 0.63 | | | |
| | Trust | 2.73 | 0.72 | 0.87** | | |
| | Perceived usefulness | 2.76 | 0.69 | 0.89** | 0.96** | |
| | Anxiety | 2.60 | 0.69 | -0.46 | -0.55** | -0.56** |

Scale is 1 (=not at all) to 5, (=very much)

* Significant at the 0.05 level, (two-tailed)

** Significant at the 0.01 level, (two-tailed)

Table 4
Coding scheme for the analysis of interview data

| Code | Definition |
|----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Aesthetics/design features | Includes comments on specific facial features (e.g., eyes, nose, hair) and overall facial appearance (e.g., attractive versus ugly-looking) |
| Expressiveness/personality | Includes comments on facial expressions (e.g., happy, sad, angry) and perceptions of personality (e. g, trustworthy, friendly, cute) |
| Gender | Actual or ascribed gender of the robot |
| General capability/ability | Includes perceptions of physical (e.g., strong) and cognitive capabilities (e.g., smart, intelligent) |
| Human-likeness | Perceived human-likeness or machine-likeness of the robot |
| Other reasons | Reasons that do not fit into any of the above codes |