國立臺灣大學理學院心理學研究所

碩士論文

Graduate Institute of Psychology

College of Science

National Taiwan University

Master Thesis

恐怖谷理論再探：機器人外觀如何影響不同年齡層使用者的喜好、個性判斷、及服務或陪伴功能的接受度

The Uncanny Valley Revisited: Age-Related Differences and the Effect of Function Type

杜昀宸

Yun-Chen Tu

指導教授：葉素玲 博士

Advisor: Su-Ling Yeh, Ph.D.

中華民國 108 年 5 月

May 2019

doi:10.6342/NTU201900877

國立台灣大學理學院心理學研究所

論文口試委員會審定書

杜昀宸先生所提論文 The Uncanny Valley Revisited: Age-Related Differences and the Effect of Function Type.

經本委員會審議，符合碩士學位標準，特此證明。

論文考試委員會

主席 岳修平

委員 葉素玲 羅仕宇

賴建益 岳修平

指導教授：葉素玲

所主任：鄭伯壎

中華民國108年4月30日

致謝

碩士生涯說短不短、說長不長，一下子也就要結束了，很感謝父母一直以來的心理跟經濟上的支持，在我做出所有選擇時都沒有太多干涉，也讓我在這段在學時間享有舒適的學習環境，而這份論文也有很大部分也實質地仰賴父母始得完成。感謝指導教授葉素玲老師提供的機會與協助，因緣際會下才有機會接觸到這個相對和實驗室原先主題關聯性不高的研究，而在問卷製作的過程、資料整理、和論文寫作的階段也和老師有多次討論，最終持著這些協助完成整份論文。感謝博士後研究員簡頌恩的協助，為了讓我的論文更趨完善，花了許多心思與時間討論實驗的設計，甚至到最後論文語句修改的階段，也都給予相當多的幫助。感謝賴岳益經理、羅仕宇教授、及岳修平教授，願意撥冗參與口試，且提供許多豐富的建議。感謝EPA實驗室的成員，實驗初期的問卷測試、實驗時的受試者招募都仰賴實驗室大家的幫忙。也謝謝在這段時間遇見的你，因為這段緣份改變我許多對生涯目標的想法，各自努力、各自加油。

doi:10.6342/NTU201900877

摘要

隨著人口出生率下降和老齡化加劇，解決照護長者勞動力短缺的問題日趨嚴峻。值此智慧製造的工業 4.0 時代，機器人可以成為有效協助長者日常生活的替代方案。然而，先前研究顯示長者對機器人的接受程度較年輕人低，推測其中一個原因是機器人的外觀所致。恐怖谷理論指的是當機器人變得更像人類時，人們會有更正向的評價，但這個趨勢只能達到某個程度；當機器人和人類外觀達到高度相像時，人們的喜好程度會急遽下降，形成一個恐怖谷的曲線。基於過去支持恐怖谷的結果主要來自年輕人，未必能適用於各年齡層，本研究檢驗恐怖谷是否也適用於老年人和中年人。我們也檢驗人們對不同功能的機器人（陪伴型或服務型）的接受度是否會根據機器人的外觀而改變，以及機器人的個性是否與其功能的接受度有關。我們讓年輕人（N = 80，年齡 18-39 歲)、中年人（N = 87，年齡 40-59 歲)、和長者（N = 88，年齡 60-87 歲）一次看一組從總數 84 張機器人的照片中選出的單一張機器人臉，並對機器人照片進行印象的評估以及關於機器人的使用意圖。結果顯示年輕人和中年人出現恐怖谷，然而老年人沒有表現出此效果—無論機器人提供哪種功能，他們都偏好由外型像人的機器人提供。除權威性外，依機器人外觀所評估的個性分數與功能接受度均呈現正相關。這些發現意味著機器人的設計應該依據不同年齡的使用者及其提供的服務類型改變外觀設計。

關鍵字： 恐怖谷理論、長者、接受度、服務型、陪伴型、機器人、外表、個性、使用者、年齡

doi:10.6342/NTU201900877

The Uncanny Valley Revisited:

Age-Related Difference and the Effect of Function Type

Yun-Chen Tu

Abstract

Due to declined birthrate and the increasing aging population, shortage of caregiving labor force has become a critical issue worldwide. Introducing robotic products could provide an effective way to help older adults' daily lives. However, previous studies indicated that older adults' acceptance of robots was lower than younger adults. One possible reason of this lower acceptance of robots might be robot appearance. The Uncanny Valley Phenomenon (UVP) refers to the phenomenon that people rate more positively as robots become more humanlike, but only up to a certain point; as it approaches near-perfect similarity of human appearance, likeability drops and forms an uncanny valley. Nonetheless, evidence for the UVP were mainly from younger adults. We therefore examined whether the UVP is also applicable for older and middle-aged adults in the present study. We also examined whether the acceptance of functions (companion vs. service) would change based on robot appearance, and whether perceived personalities have any relation with the acceptance of robot

doi:10.6342/NTU201900877

function. We asked younger (*N*= 80, age 18-39), middle-aged (*N*= 87, age 40-59), and older (*N*= 88, age 60-87) adults to view each picture of a set of robot pictures selected from a totally 84 robots and evaluate their impression on each robot and the intention of use regarding robot functions. UVP was found in younger and middle-aged adults; however, older adults did not show UVP. They preferred humanlike over non-humanlike robots, regardless of robot function. Scores on each personality－except for authoritativeness－showed positive correlations with the acceptance of functions. These findings suggest that the design of assistive robots should take UVP into consideration by customizing robot appearance in accordance with the function provided to different age groups.

Keywords: uncanny valley, age, older adults, acceptance, appearance, personality, robot, companionship, service function

doi:10.6342/NTU201900877

Table of Contents

doi:10.6342/NTU201900877

doi:10.6342/NTU201900877

List of Tables and Figures

doi:10.6342/NTU201900877

doi:10.6342/NTU201900877

doi:10.6342/NTU201900877

Introduction

Due to low birth and death rate, there is a global trend of aging population, resulting in shortage of caregivers for older adults (Ministry of the Interior, 2018; World Health Organization, 2011). Studies have shown that there is increasing proportion of older adults who live alone across different countries (Macunovich, Easterlin, Schaeffer, & Crimmins, 1995; Reher & Requena, 2018; Tomassini, Glaser, Wolf, van Groenou, & Grundy, 2004). Older adults living alone would easily feel loneliness, which is linked to both mental and physical problems (Yeh & Lo, 2004). Also, as people age, they have higher possibility of falls, hospitalizations, disability, and death (Fried et al., 2001). Despite that older adults tend to have higher need of home cares, however, their offspring might not be able to take care of them due to various reasons. Hiring caregivers to help older adults in their daily lives could be one way to solve the problem, but evidently the labor force of caregivers is not enough for everyone. Alternatively, assistive robots may serve as an efficient solution for the shortage of labor in the aging population (Roy et al., 2000; Smarr, Fausset, & Rogers, 2011).

However, previous studies have shown that older adults had more negative attitude toward robots than their younger counterpart. Chien et al. (2019) used implicit

doi:10.6342/NTU201900877

association task to examine the negative attitude toward robots and results showed that compared to younger adults, older adults had stronger association between negative words and robots, as well as lower curiosity toward robots. And users' attitudes toward robots would affect their acceptance of robotic products (Heerink, Kröse, Evers, & Wielinga, 2010; Nomura, Kanda, Suzuki, & Kato, 2008). Understanding how to enhance older adults' acceptance of robots is thus critical, otherwise even when robots are designed with accurate and fine functions, older adults would still refuse to use them.

Indeed, understanding user acceptance of robots is a critical component of designing robots. Beer, Prakash, Mitzner, and Rogers (2011) pointed out three factors that would affect user acceptance of robots: appearance, functionality, and social intelligence. First of all, appearance would change the acceptance. For example, Robots' human-likeness, structure, and gender could all possibly change users' acceptance. Secondly, functionality includes sub-factors of task types, autonomy, and control interface. Users prefer robots that could offer assistance they need. Users also prefer robots that can overcome the variation in the environment. For example, current robot vacuums can only operate on a flat ground, but users would expect that the robot could move between different floors by themselves. Users also care about the control interface. They prefer voice control rather than manual manipulation. Last,

doi:10.6342/NTU201900877

users also prefer robots with social ability, which enables robots to express emotion, show non-verbal emotional cues, and elicit users to perform social behaviors to them mindlessly. Such robots allow users to project their thoughts on robots and categorize robots with specific personality. Also, if robots could express emotions, users might interact with robots who could express emotions differently compared to robots that could not. Users prefer robots can follow their eye movements (implicit non-verbal cue) or nod when agreeing with them (explicit non-verbal cue).

Several research demonstrated that appearance is one crucial factor that change the user acceptance of robots (Bartneck, Kanda, Ishiguro, & Hagita, 2007; Mathur & Reichling, 2016; Mori, 1970; Seyama & Nagayama, 2007). Mori (1970) first introduced the Uncanny Valley Phenomenon (UVP), which refers to the phenomenon that people rate more familiar as robots become more humanlike, but only up to a certain point; as it approaches near-perfect similarity of human appearance, familiarity drops and forms an uncanny valley. Because of differences in translation, some researchers also used other positive descriptions such as likeability, affinity, and pleasantness instead of familiarity as question statements for participants to rate, and similar effect was found still (Mathur & Reichling, 2016; Seyama & Nagayama, 2007).

Nevertheless, evidence supporting the existence of UVP were mainly from

doi:10.6342/NTU201900877

younger adults whose age were lower than 40 (Bartneck et al., 2007; Ho & MacDorman, 2017; Seyama & Nagayama, 2007). Other studies recruited infants and teenagers as participants to examine the minimum age to yield the UVP (Lewkowicz & Ghazanfar, 2012; Tung, 2016). These studies showed that the UVP might change across different age groups. For example, Lewkowicz and Ghazanfar (2012) found that infant did not show the UVP until one year old, suggesting that the UVP is not inherent, but learnt from life experience. However, no study so far has explored whether the UVP would change after age 40. Even some studies (Mathur & Reichling, 2016) included participants over age 40, data from all age groups were mixed together instead of separated them based on age. Thus, it is important to examine whether the UVP could still be observed for older adults, especially when designing robots for older adults at aging societies are imperative as mentioned above.

We predict that older adults would show weakened UVP than younger adults for the following reason. Previous studies have revealed the relations between the UVP and three theoretically motivated trait indices: religious fundamentalism, animal reminder sensitivity, and trait anxiety (MacDorman & Entezari, 2015). First, for fundamentalists, human beings are unique and different from other creatures. Therefore, people with high tendency on this trait will intensify the drop in the UVP (Vail et al., 2010; Vess, Arndt, & Cox, 2012). Second, the animal reminder sensitivity

doi:10.6342/NTU201900877

hypothesis places much emphasis on the difference between living and non-living things. People perceive some robots as eerier because they are disturbed by those robots with both nonliving and living features. Then, these uncertain or conflictory cues induce negative feelings for the observers. Last but not least, trait anxiety indicates that people with higher anxiety feel that robots are more unpredictable, so they perceive robots as more aversive than people with lower anxiety (Samochowiec & Florack, 2010). Among these three factors, the degree of animal reminder sensitivity and anxiety were inversely correlated with age, and thus older adults might have lower UVP due to their lower animal reminder sensitivity and anxiety.

In addition to examining what kind of robot appearance older users would prefer, we also investigated whether appearance would affect their acceptance of robots' assistance. We predict that users tend to accept assistance from robots with higher likeability. Chu et al. (2019) used a semi-structural questionnaire to interview older adults, asking them to draw and describe their ideal robots. Also, they introduced one companionship robot (PARO) and one service robot (Zenbo) to older adults, and let older adults fill out questionnaires about acceptance of robots. Results demonstrated that older adults had more positive attitude toward the service robot than the companion one, and they perceived the service robot as more useful. Older adults highly anticipated that robots should be able to do housework; that is, a robot with

doi:10.6342/NTU201900877

service functions, including finding or fetching items, getting outside (as a tool), making meal, and so on. With a companion robot, older adults expected robots to do things like chatting, staying close by, and entertainment. Ezer, Fisk, and Rogers (2009) also found that older adults needed robots to help them do chores, ensure security, provide physical aid, cook, and maintain or repair things. Although these studies found that older adults tended to accept service function than companionship from robots, the companionship is still important. For example, Yeh and Lo (2004) found older adults who perceived the feeling of loneliness were easily linked to both mental and physical problems, and robots could offer companionship to older adults to reduce the negative effects from loneliness. Wada and Shibata (2007) placed animal-like companion robot, PARO, in the care house, and they found that older adults obtained both socio-psychological and physiological benefits from the presence of the companion robot. As a result, in the current study we examined the relation between appearance and acceptance of not only service function but also companionship.

Apart from the relation between appearance and acceptance of functions, it is also possible that different appearances of robots would also induce different perceived personalities and consequently affect the acceptance of robots, as previous studies showed that appearance strongly influences interpersonal relationships.

doi:10.6342/NTU201900877

Albright, Kenny, and Malloy (1988) found that people could correctly perceive personalities of strangers based on first impressions. Mount, Barrick, and Stewart (1998) demonstrated that among the Five-Factors Model of personality, Conscientiousness, Agreeableness, and Emotional Stability are positively related to performance in jobs. Other studies also found that employee's personalities would affect subjective rating scores on work performance from their supervisors (Barrick & Mount, 1991; Brown, Mowen, Donavan, & Licata, 2002). In addition, we included the rating of authoritativeness in the present study. Authoritativeness refers to whether one's opinion would be accepted because of his or her expertise (McCroskey, 1966). According to Lee, Peng, Jin, and Yan (2006), users could perceive robots' personalities and change their interaction styles with robots accordingly. Thus, we also examined whether perceived personality could affect the acceptance of companionship/service function. We predicted that users would be more likely to be accompanied and served by robots with conscientiousness, agreeableness, and emotional stability. However, we predicted that people would like to be served by robots with authoritativeness because people would be willing to be served by professional workers. And people would not change their acceptance of companionship based on authoritativeness because they care whether robots could offer them social supports instead of professional advices.

doi:10.6342/NTU201900877

To sum up, we found that designs for robot could affect users acceptance. In the present study, we aimed to enhance older adults' acceptance of robots. Thus, we had three main objects: (1) to examine whether the uncanny valley phenomenon varied across different age groups, (2) to examine whether the acceptance of companion and/or service functions would change based on robot appearance, and what would be an ideal design of robot appearance according to the function which the robot could offer, (3) to investigate whether perceived personalities of robots had any relation with the acceptance of companionship and service functions.

In the present study, we aimed to examine the relation between acceptance and appearance not only for older adults but also for middle-aged adults (40 – 59 years old) because of the following reasons. First, middle-aged adults will also need robots' assistance in their near future because people now live longer than before in the era of labor shortage of caregivers. Second, compared to older adults, middle-aged adults may have higher possibility to access information of robots, and thus they may have higher acceptance of robots' assistance. Third, currently, middle-aged adults tend to realize that their children might not be able to take care of them after they get older, so they would be more willing to accept assistances from robots.

We also included younger adults as a basis for comparison. Because previous results of UVP were mainly from younger adults, and it is imperative to see whether

doi:10.6342/NTU201900877

our results could replicate previous studies. Furthermore, to design the appearance of robots for all age groups, it is important to see the overall picture by comparing the three age groups and see whether there is age differences in the UVP, the relationship between acceptance and appearance of robots in terms of different function types, and the relationship between the perceived personality from appearance and the acceptance of different functions.

doi:10.6342/NTU201900877

General Method

Participants

We recruited 255 participants in total, including 80 young adults (aged from 18-39, average = 21.87, male 45%), 87 middle-aged adults (aged from 40-59, average = 50.26, male 26%), and 88 old adults (aged above 60, average = 65.18, male 44%). Participants were recruited from websites, universities, and hospitals. Participants received NT$ 140 or course credit.

Questionnaire

We collected data via an on-line questionnaire created on SurveyMonkey (https://www.surveymonkey.com/), after participants gave their consent and practiced to answer three types (multiple choices, slider, and blank filling) of questions to ensure that participants from all three age groups understand how to respond to the on-line questionnaire. Participants were required to complete three questionnaires: (1) State Curiosity questionnaire, (2) Rating of uncanny valley, personalities, and acceptance of functions, (3) Questionnaire of participants' demographic information. Participants need to watch robot faces and rate on a series of questions in the second part. Before participants started to evaluate each robot, they were asked to carefully

doi:10.6342/NTU201900877

check all 84 pictures of robots and establish the best and worst benchmark of likeability among them.

State Curiosity questionnaire. State curiosity toward robots, which was revised from its original version targeting to measure participants' curiosity toward a specific kind of sports (Park, Mahony, Kim, & Do Kim, 2015), was used to measure participants' curiosity toward the robot. The questionnaire contains 6 items (see Table 1). Participants gave their rating scores via a 7-point Likert scale (appendix A.)

Table 1

State curiosity questionnaire

Questionnaire	**Items**
State curiosity toward robots	1. "How curious do you feel about this robot?" 2. "How likely would you spend time watching a video that introduces this robot?" 3. "How much do you want to know about this robot?" 4. "How likely would you read the information about the robot if you have the instruction about using the robot?" 5. "How likely would you actively find the information related to the robot?" 6. "How likely would you actively participate at a course related to the robot use if the course is held nearby your home?"

doi:10.6342/NTU201900877

Rating of uncanny valley, personalities and acceptance of functions. Seventeen questions were included in this part, along with a robot face for each question. Based on Chu et al. (2019), we set up four questions for service functions and four questions for companionship by combining the functions obtained in Chu et al. (2019) that had the highest scores, meaning that older adults care most about what robots could help them. Service functions included those of daily life, protection, touch, and privacy, and companionship functions included those of staying close by, entertainment, chatting, and sharing feelings. Participants were required to use the slider range from -100 (strongly disagree) to +100 (strongly agree) to rate their willingness to use these functions.

After the eight questions about the two types of functions, four questions related to perceived personalities of the robot were also shown in the questionnaire (trustworthiness/reliability, friendliness/approachability, adorability, and authoritativeness, which were related to willingness of use). Furthermore, three commonly used questions in the UVP study, such as likeability, disgust, and humanness were also presented and recorded. Finally, an exclusion question was placed to avoid the effect from prior experience. Participants needed to check whether they had seen the robot before via choosing one answer among "Yes", "No", and "Uncertain". Only those data from a given robot face which participants never saw it

doi:10.6342/NTU201900877

before would be analyzed. In this part, 16 questions were recorded by a slider ranged from -100 to 100, and one question was recorded by multiple choices (appendix B.)

Table 2

Rating of uncanny valley, personalities, and acceptance of function

Question type	Index	Item
Service function	Daily life	1. When I can' t take care of myself, I am willing to have this robot assist me with my daily life (including housekeeping, cleaning, meal preparation, and delivery).
	Protection	2. I am willing to let this robot protect me.
	Touch	3. I am willing to be touched by this robot (including assist in moving, massage, and hold up, etc).
	Privacy	4. I am willing to have this robot assist me in dealing with things involving personal privacy (including bathing, dressing, feeding, and financial management).
Companion function	Company	5. I am willing to let this robot accompany me.
	Entertainment	6. I am willing to spend my leisure time and have fun with this robot.
	Chatting	7. I can chat with this robot as if it were a friend of mine.
	Sharing feelings	8. I am willing to share my thoughts and feelings with this robot.
Uncanny valley	Likeability	9. I _____ this robot. dislike – like
	Disgust	10. I feel _____ by this robot. strongly non-disgusted – strongly disgusted

doi:10.6342/NTU201900877

Table 3

Rating of uncanny valley, personalities, and acceptance of function (continued)

Question type	Index	Item
Personality	Reliability and trustworthiness	11. This robot looks _____ very unreliable and not trustworthy – very reliable and trustworthy
	Friendly and approachable	12. This robot looks _____ very unfriendly and unapproachable – very friendly and approachable
	Authoritativeness	13. This robot looks _____ very unauthoritative – very authoritative
	Adorability	14. This robot looks _____ not very adorable – very adorable
	Sex	15. What gender is this robot? Male – uncertain - female
Uncanny valley	Humanness	16. This robot ______ human. doesn't look like – does look like
Exclusion question		17. I have seen this robot before. (Yes / No / uncertain)

Questionnaire of participants' demographic information. We also collected participants' demographic information such as age, sex, and education level. Also, they needed to answer their willingness to buy a robot because we aimed to understand how likely the potential clients would like to spend money on the product. Besides, they were asked to subjectively and objectively estimate their usage time of electronic products. We also asked whether participants had used robots in the past,

doi:10.6342/NTU201900877

and how frequently they used robots.

Stimuli

Eighty-four pictures of robot face were adopted as stimuli. We mainly used the 79 pictures of robots (Figure 1) from Mathur and Reichling (2016). One picture was excluded because the robot was designed based on one professor's appearance and we wanted to exclude the effect from participants' experience. Details could be found in Mathur and Reichling (2016). We also added four additional robots (Alpha One Pro, I-Cat, Paro, and Robohon) to include the robots which were recently designed. Google image in the internet was used to find all pictures based on the specific keywords such as "robot face", "interactive robot", "human robot", and "robot". All pictures included meet the following inclusion and exclusion criteria.

Inclusion criteria

1. Full face is shown from top of head to chin.

2. Face is shown in frontal to 3/4 aspect (both eyes visible).

3. The robot is intended to interact socially with humans.

4. The robot has actually been built.

5. The robot is capable of physical movement (e.g., not a sculpture or purely computer generated image, CGI, representation that lacks a three-dimensional body structure).

6. The robot is shown as it is meant to interact with users (e.g., not missing any hair,

doi:10.6342/NTU201900877

facial parts, skin, or clothing, if these elements are intended.

Exclusion criteria:

1. The image includes other faces or human body parts that would appear in the final cropped image.

2. Objects or text overlap the face.

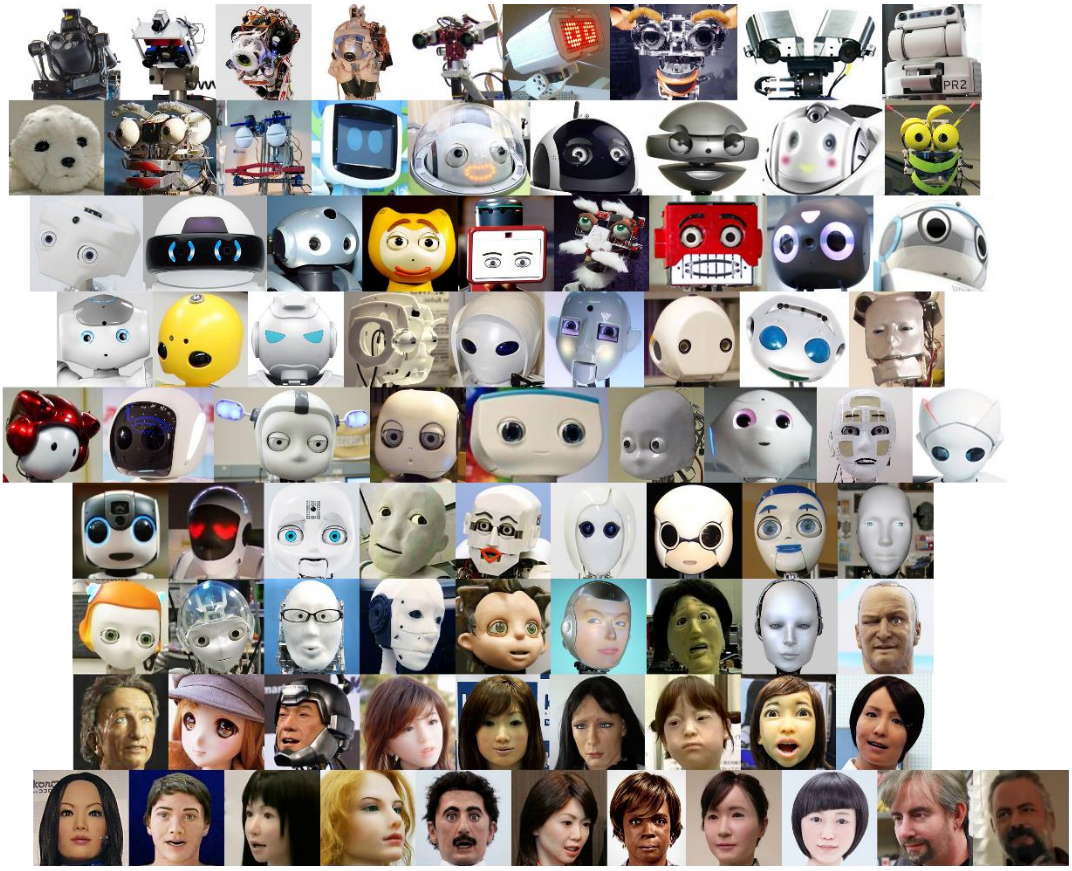

Figure 1. Robots picture (Non-humanlike to humanlike from top to bottom panel).

Procedure

In the beginning of the questionnaire, they were informed that their data might be used as references for design guidelines for actual robotic products and they were asked to carefully look at pictures and rated scores for all questions. If they failed to

doi:10.6342/NTU201900877

correctly respond to a specific question (designed to ensure their paying attention to the questionnaires and their answers), they could not get the reward for their participation. In order to ensure that they focused on the whole process and to keep them from feeling tired, each participant only rated 16 robots, which contained 15 robots randomly chosen from 83 possible pictures of robots and one picture for checking whether they had concentrated on the given task. In addition, one question was included to check whether they stayed focus on the task. If they did not follow the instruction to give +43 score on the checking question "I am willing to let this robot accompany me," the task would be ended and data from the participant who failed to fill the requirement would be abandoned.

Data analysis

We used Python to transform from the original data to the analysis data, and used R to complete all analyses. All data were merged into 83 data points for each picture (excluding the checking question picture). Because all participants would score same questions for different pictures of robots, we needed to ensure that the responses to the questions were independent from each other. Thus, the Intra-Class Correlation (ICC) was used to examine whether all participants showed similar tendency on rating specific questions. After ensuring that there was no such tendency, we used these points for modeling the relationship between the rating score of each function and the

doi:10.6342/NTU201900877

score of personalities or humanness from four-degree model to zero-degree model via polynomial regression. Each model was examined by the analysis of variance (ANOVA) framework to compare the variability of rating scores between and within each picture, and to find the best model to explain the data. Stepwise regression was used to compare all suitable models from the most complicated model to the least complicated one. Based on results from the stepwise regression analysis, we selected the least complicated model with high interpretation for all data. We separated data into three age groups: younger, middle-aged, and older adults.

doi:10.6342/NTU201900877

Results

The uncanny valley across different age groups

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0, 0.07), indicating that each robot got different scores on humanness. Thus, participants were able to categorized our stimuli from non-humanlike to humanlike robots by participants. The result of the robots' likeability score also had low-within subject clustering (ICC= 0.09; 95% CI: 0.05, 0.15). Among all model, the third-degree model was the most suitable model to explain the results (third-versus-second: $F(1, 79) = 21.341$, $p < .01$; fourth-versus-third: $F(1, 78) = 0.409$, $p = .524$). The model showed a classic pattern of UVP. According to this model (Figure 2, red line, $R^2_{adj} = 0.324$), when the humanness score reached 50.24, the likeability reached the lowest turning point, -38.62, in the rating ranged between -100 and 100. On the contrary, when the humanness score reached -49.85, the likeability reached the highest turning point, 20.45. Figure 3 depict the top and bottom five preferable appearances for younger adults and it is evident that younger adults preferred non-humanlike robots to human-like ones.

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low

doi:10.6342/NTU201900877

within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The rating results of the robots' likeability score also had low-within subject clustering (ICC= 0.22; 95% CI: 0.16, 0.30). Among all model, the third-degree model best interprets the data (third-versus-second: $F(1, 79) = 10.918$, $p = .001$; fourth-versus-third: $F(1, 78) = 0.728$, $p = .396$), which also showed a UVP pattern (Figure 2, green line, $R^2_{adj} = 0.134$). According to this model, when the humanness score reached 42.51, the likeability score reached its lower turning point, -20.28. On the contrary, when the humanness score reached 3.86, the likeability score reached its higher turning point, -28.03. Figure 4 showed the top and bottom five preferable appearances of robots for middle-aged adults. It can be seen that middle-aged adults had lower preference for non-humanlike robots than younger adults.

Older adults. The result from older adults also showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The rating result on the robots' likeability score also had low-within subject clustering (ICC= 0.22; 95% CI: 0.16, 0.29). Among all models, the first-degree model best interprets the data (first-versus-no: $F(1, 81) = 10.686$, $p < .01$; second-versus-first: $F(1, 80) = 3.364$, $p = .07$; third-versus-second: $F(1, 79) = 1.446$, $p = .233$; fourth-versus-third: $F(1, 78) = 0.399$, $p = .53$), suggesting that likeability score was proportional to robot's humanness

doi:10.6342/NTU201900877

(Figure 2, blue line, $R^2_{adj} = 0.103$). That is, no UVP was observed in the old-adults group. Figure 5 shows the top and bottom five preferable appearances of robots for older adults, which can be seen that older adults had higher preference for humanlike than non-humanlike robots.

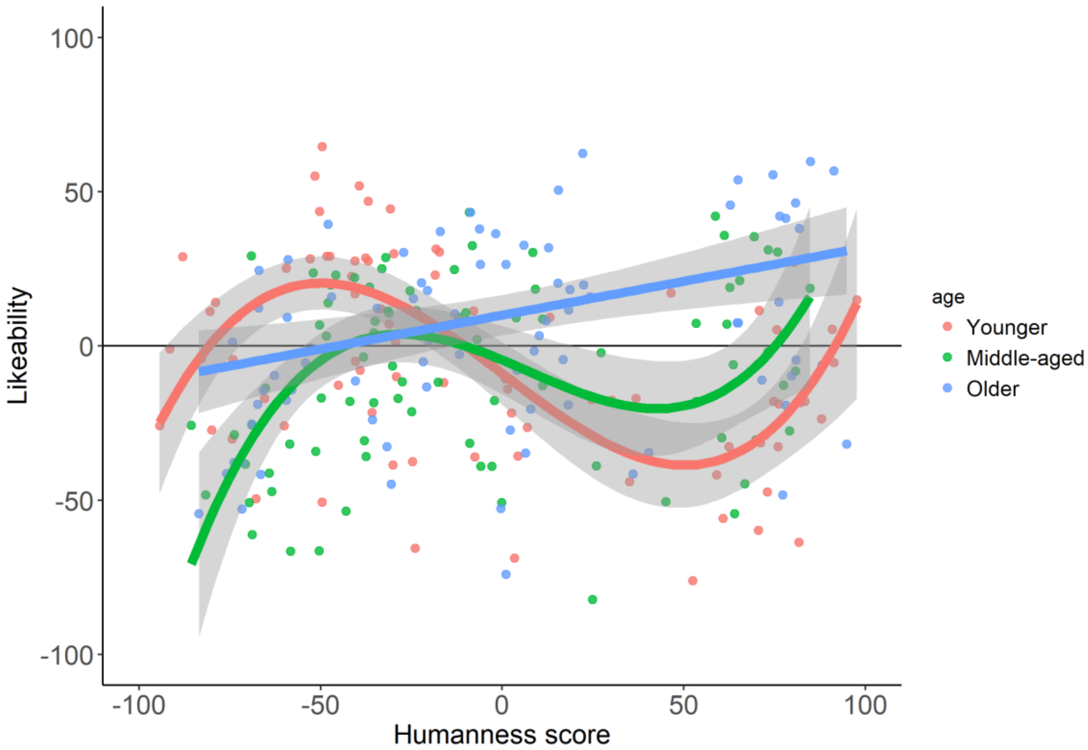

Figure 2. The relation between likeability and humanness across different ages.

doi:10.6342/NTU201900877

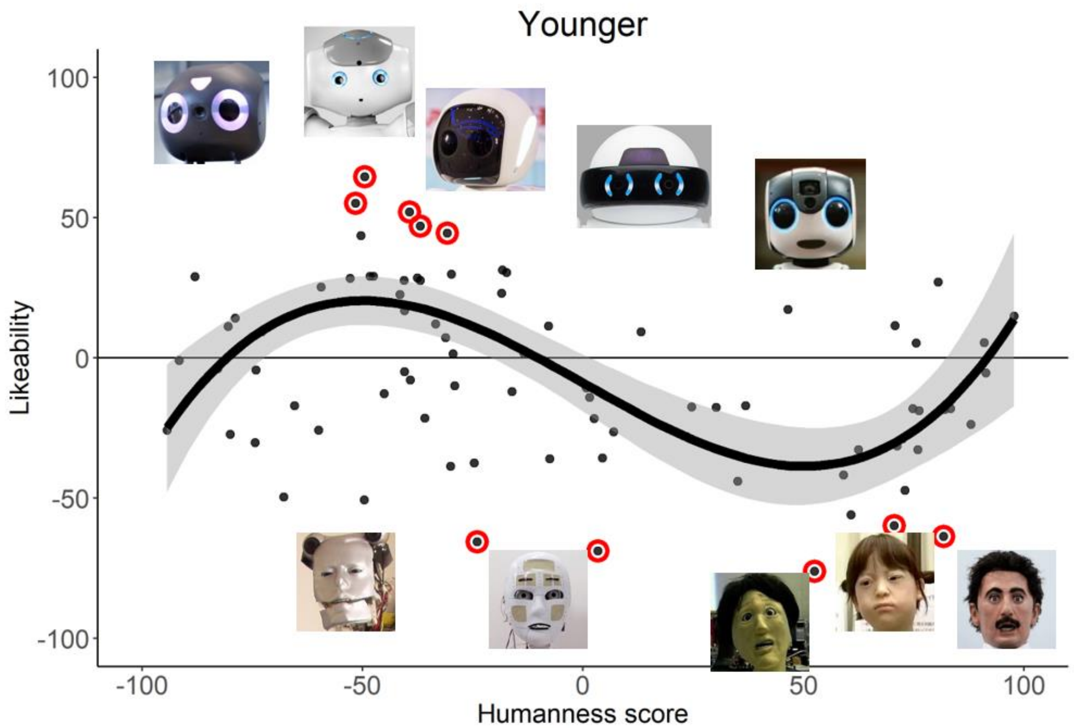

Figure 3. The top and bottom five preferable appearance of robots based on data from the younger adults group.

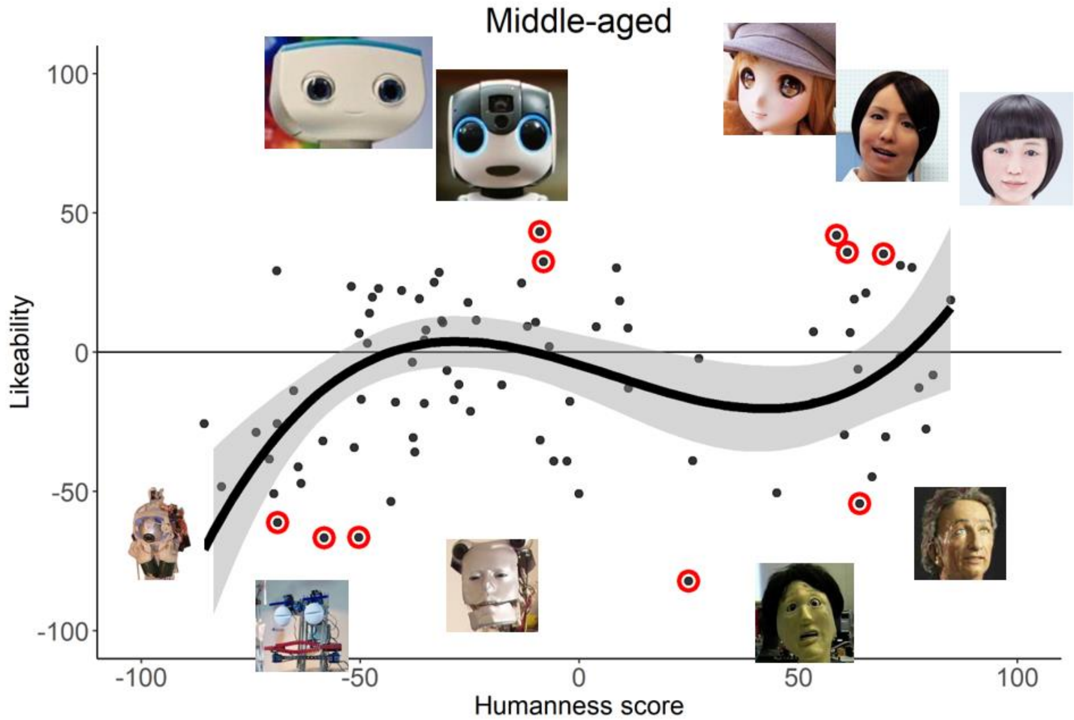

Figure 4. The top and bottom five preferable appearance of robots based on the data from middle-aged adults group.

doi:10.6342/NTU201900877

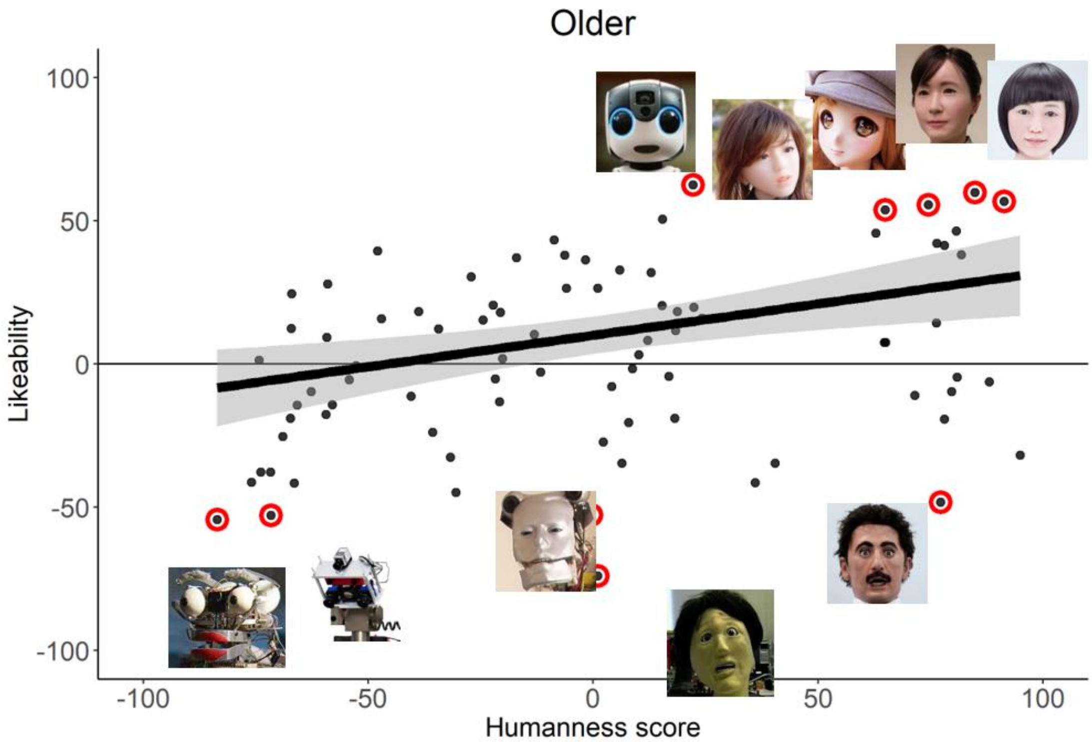

Figure 5. The top and bottom five preferable appearance of robots based on the data from older adults group.

Acceptance of functions and appearance

Companionship.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0.00, 0.07). The rating results on the acceptance of companionship also had low-within subject clustering (ICC= 0.2; 95% CI: 0.16, 0.27). Among all models, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 22.832$, $p < .01$;fourth-versus-third: $F(1, 78) = 0.156$, $p = .693$). This model showed the UVP pattern (Figure 6, red line, $R^2_{adj} = 0.271$). When the humanness score reached 46.54, the likeability reached the lowest score, -36.57.

doi:10.6342/NTU201900877

On the contrary, when the humanness score reached -50.15, the acceptance reaches the highest score, 14.23. Figure 7 shows the top and bottom five robots with companionship for younger adults. As can be seen, younger adults were more willing to be accompanied by non-humanlike robots.

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The rating results of the acceptance of companionship also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.38). Among all model, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 6.544$, $p < .05$;fourth-versus-third: $F(1, 78) = 1.927$, $p = .169$). This model showed the UVP pattern (Figure 6, green line, $R^2_{adj} = 0.131$). When the humanness score reached 41.04, the likeability reached the lowest score, -12.95. On the contrary, when the humanness score reached -0.11, the acceptance reached the higher score, -22.15. Figure 8 shows the top and bottom five robots with companionship for middle-aged adults. As can be seen, middle-aged adults were less willing to be accompanied by non-humanlike robots than younger adults did.

Older adults. The result from older adults showed that individual participants have different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The rating results on the

doi:10.6342/NTU201900877

acceptance of companionship also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.37). Among all model, the first-degree model best interprets data (first-versus-zero: $F(1, 81) = 16.533$, $p < .001$; second-versus-first: $F(1, 80) = 1.468$, $p = .229$; third-versus-second: $F(1, 79) = 2.015$, $p = .16$;fourth-versus-third: $F(1, 78) = 0.016$, $p = .901$). This model showed simple linear correlation but not the UVP (Figure 6, blue line, $R^2_{adj} = 0.158$). Figure 9 shows the top and bottom five robots with companionship for older adults. As can be seen, older adults were more willing to be accompanied by humanlike robots.

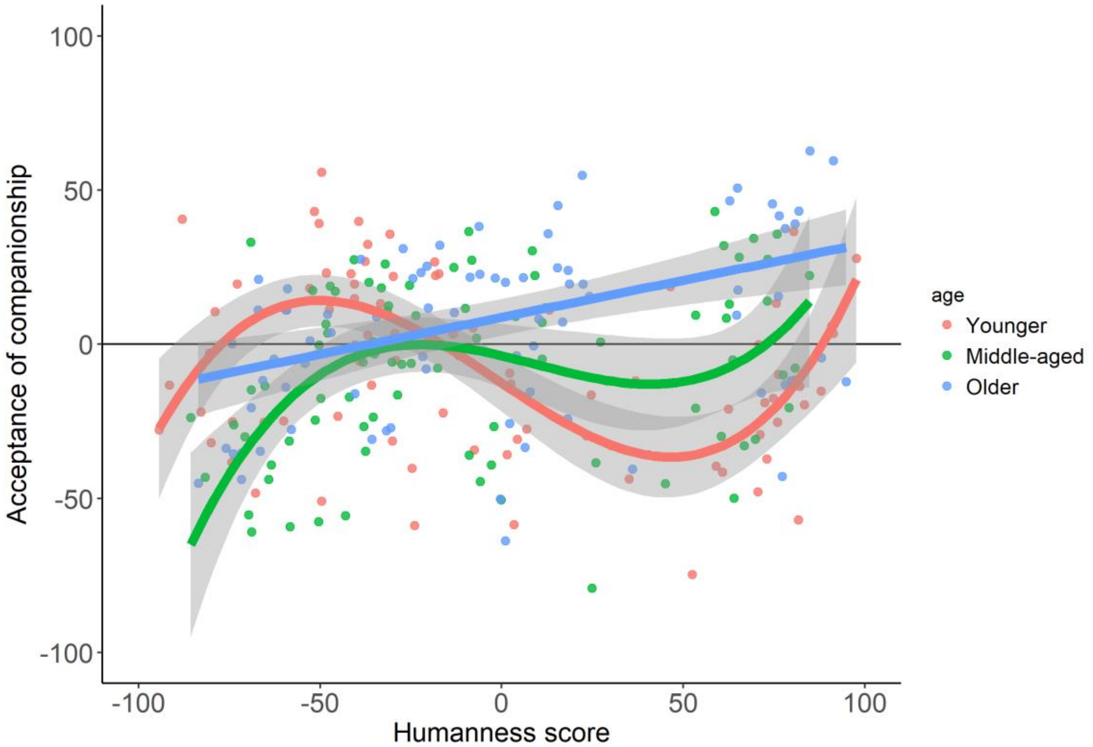

Figure 6. The relation between acceptance of companionship and humanness across different age.

doi:10.6342/NTU201900877

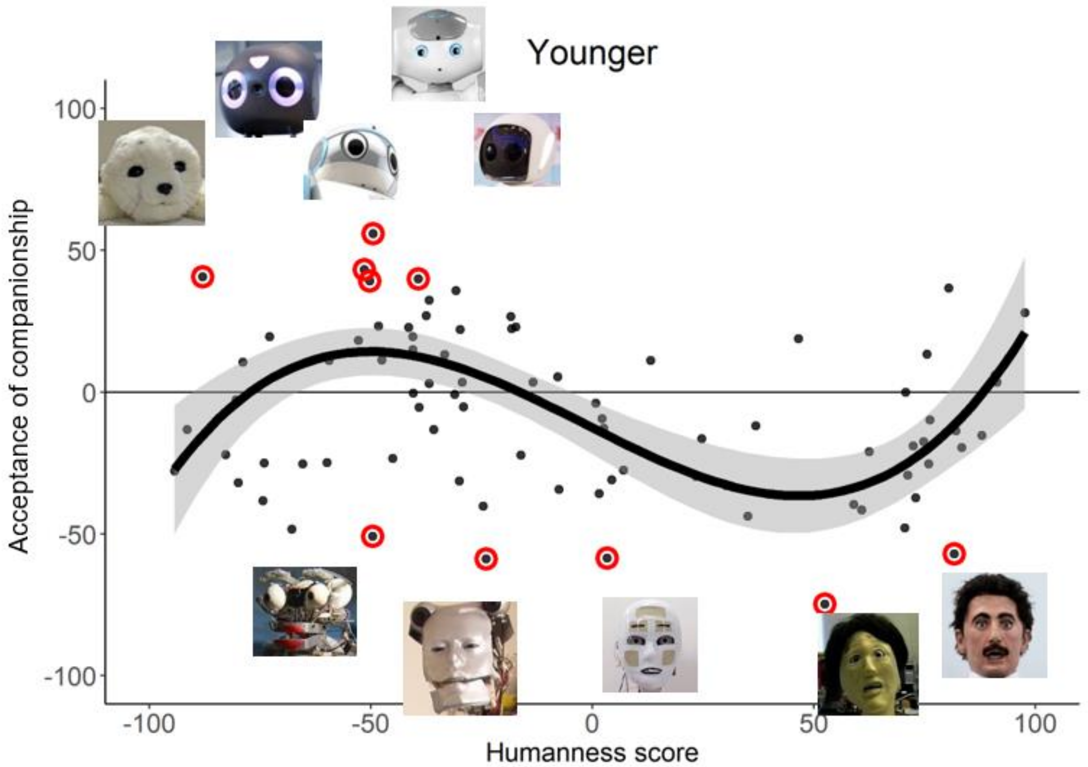

Figure 7. The top and bottom five preferable appearance for companion robots based on younger adults group.

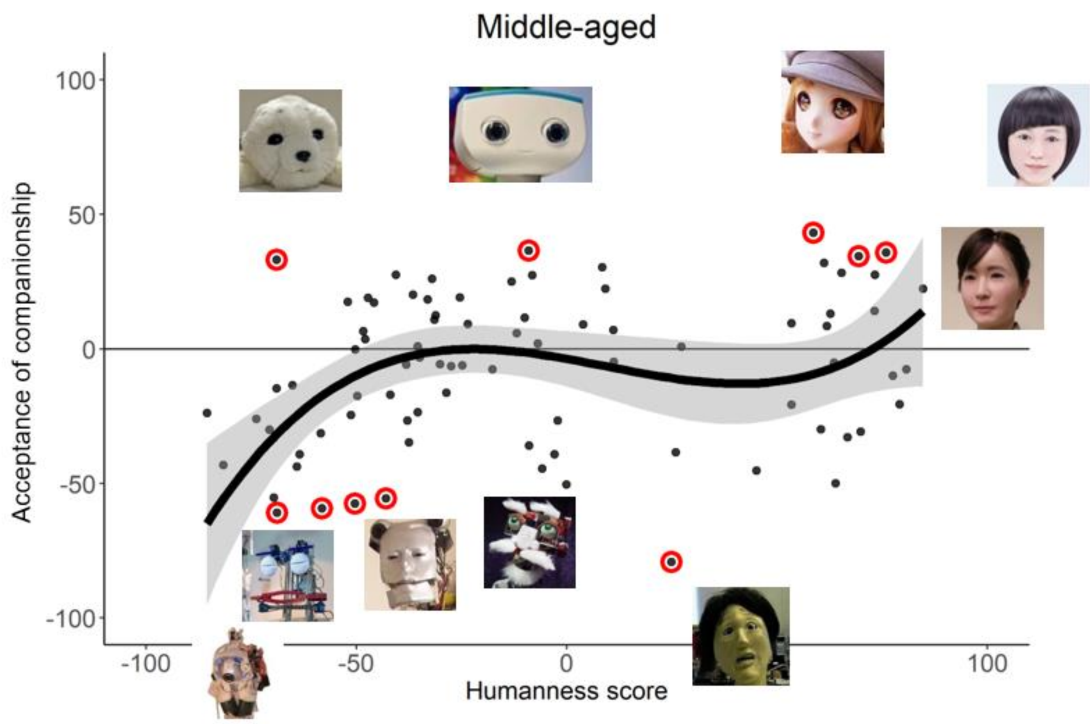

Figure 8. The top and bottom five preferable appearance for companion robots based on middle-aged adults group.

doi:10.6342/NTU201900877

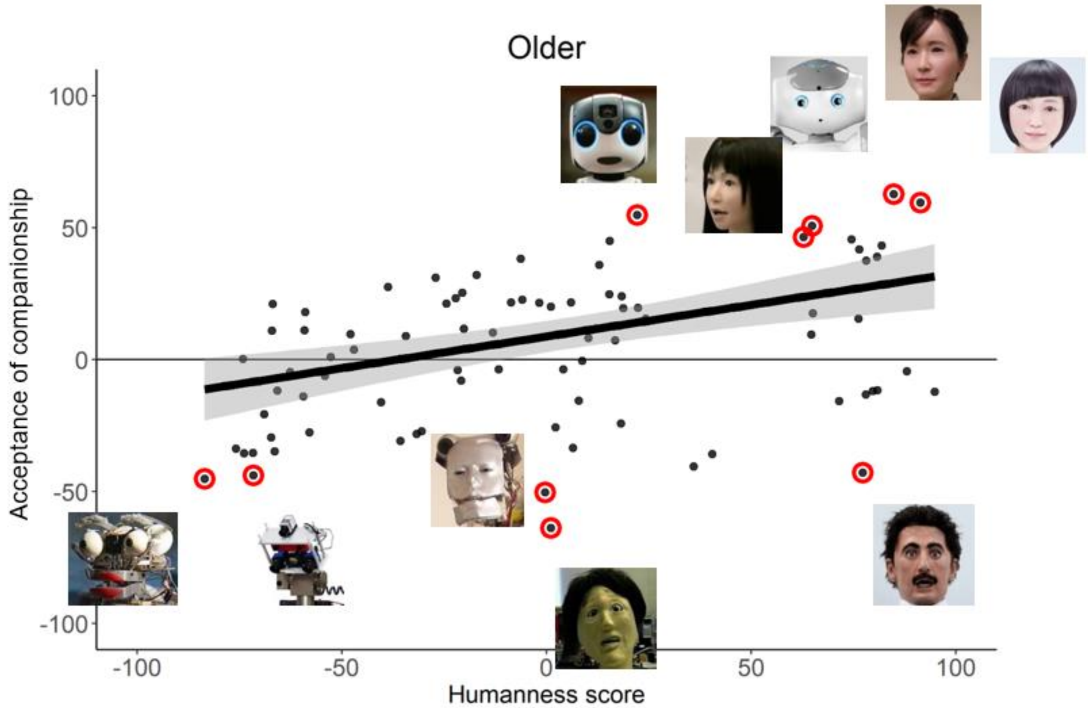

Figure 9. The top and bottom five preferable appearance for companion robots based on older adults group.

Service function.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0, 0.07). The rating results on the robots' service score also had low-within subject clustering (ICC= 0.16; 95% CI: 0.12, 0.22). Among all model, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 16.057$, $p < .001$;fourth-versus-third: $F(1, 78) = 0.732$, $p = .395$). This model showed UVP pattern (Figure 10, red line, $R^2_{adj} = 0.247$). When the humanness score reached 48.57, the acceptance of service function reached the lowest score, -32.49. On the contrary, when the humanness score reached 18.82, the

doi:10.6342/NTU201900877

acceptance reached highest score, -51.38. Figure 11 shows the top and bottom five robots with service function for younger adults. As can be seen, younger adults were more willing to be served by non-humanlike robots

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The rating result on the acceptance of service function also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.38). Among all model, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 7.272$, $p < .01$;fourth-versus-third: $F(1, 78) = 1.008$, $p = .319$). The model showed the UVP pattern (Figure 10, green line, $R^2_{adj} = 0.139$). When the humanness score reached 41.85, the acceptance of service function reached the lowest score, -13.26. On the contrary, when the humanness score reached 1.4, the acceptance reached the highest turning point, -22.17. Figure 12 shows the top and bottom five robots with service function for middle-aged adults. As can be seen, middle-aged adults were less willing to be served by non-humanlike robots than younger adults did.

Older adults. The result from older adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The rating result on the robots'

doi:10.6342/NTU201900877

service score also had low-within subject clustering (ICC= 0.28; 95% CI: 0.23, 0.36). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 14.648$, $p < .001$second-versus-first: $F(1, 80) = 1.488$, $p = .226$;third-versus-second: $F(1, 79) = 1.254$, $p = .266$;fourth-versus-third: $F(1, 78) = 0.026$, $p = .871$). The model showed no UVP pattern but a linear relationship (Figure 10, blue line, $R^2_{adj} = 0.143$). Figure 13 shows the top and bottom five robots with service function for older adults. As can be seen, older adults were more willing to be served by humanlike robots

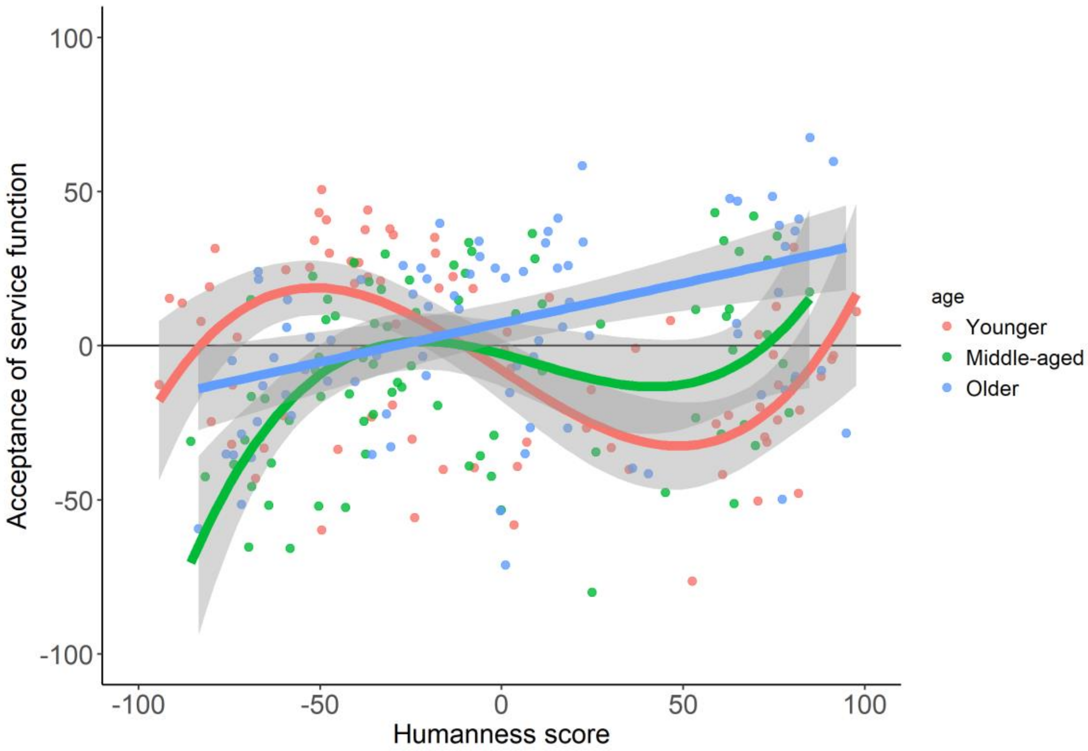

Figure 10. The relation between acceptance of service function and humanness across different age.

doi:10.6342/NTU201900877

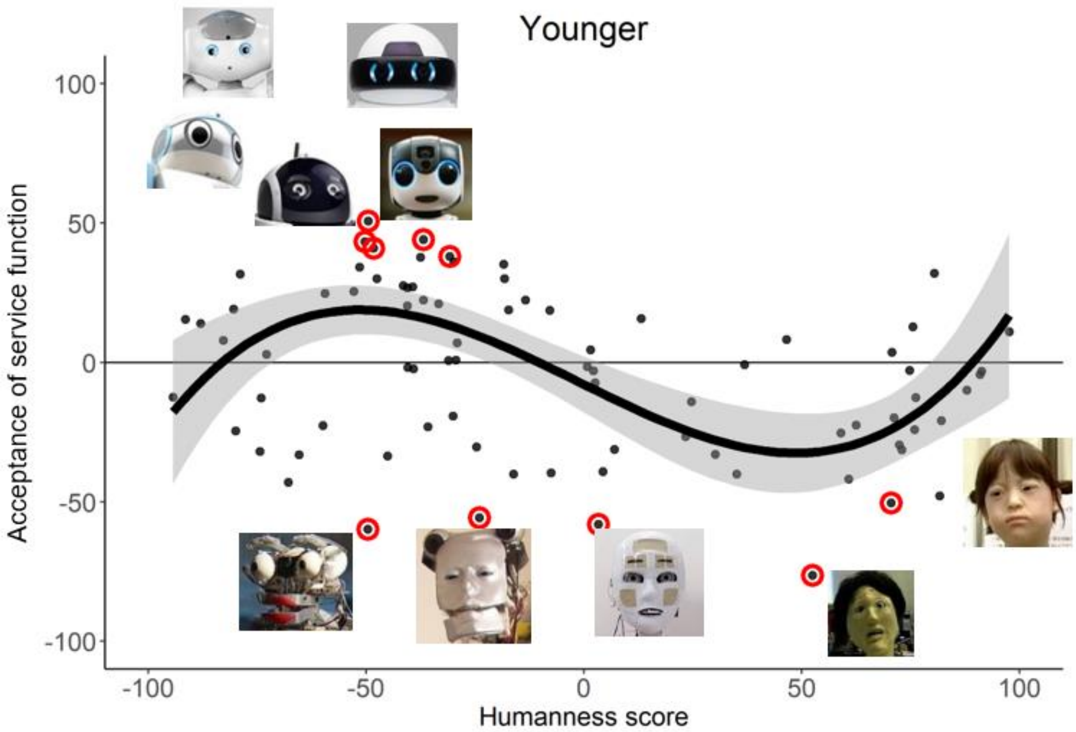

Figure 11 The top and bottom five preferable appearance for service robots based on younger adults group.

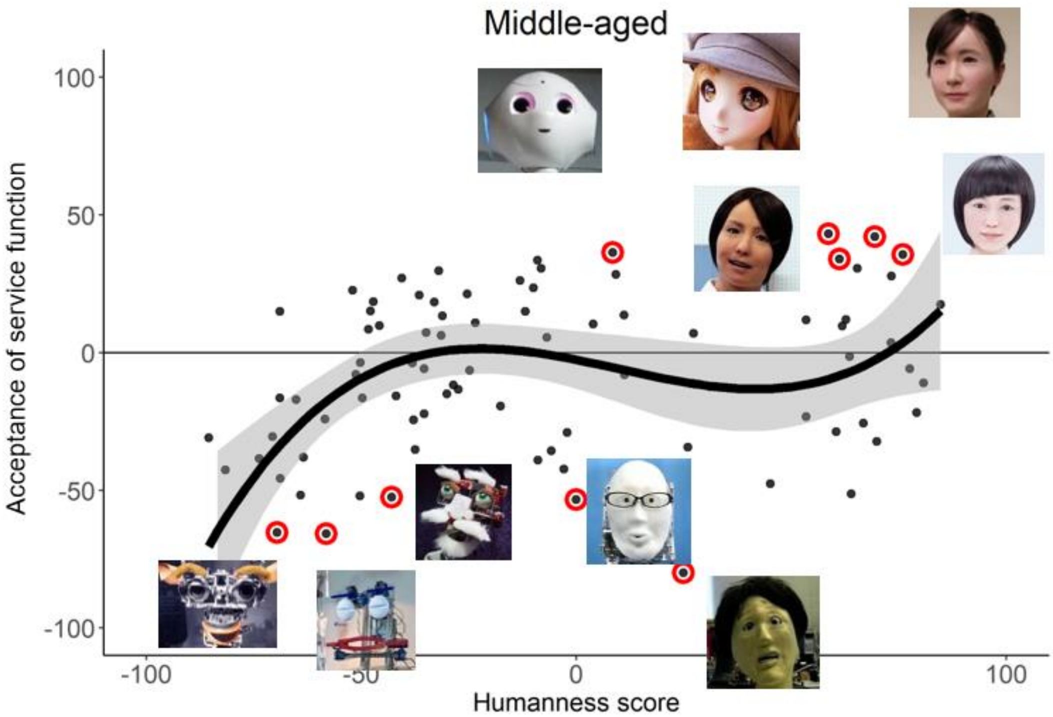

Figure 12. The top and bottom five preferable appearance for service robots based on middle-aged adults group.

doi:10.6342/NTU201900877

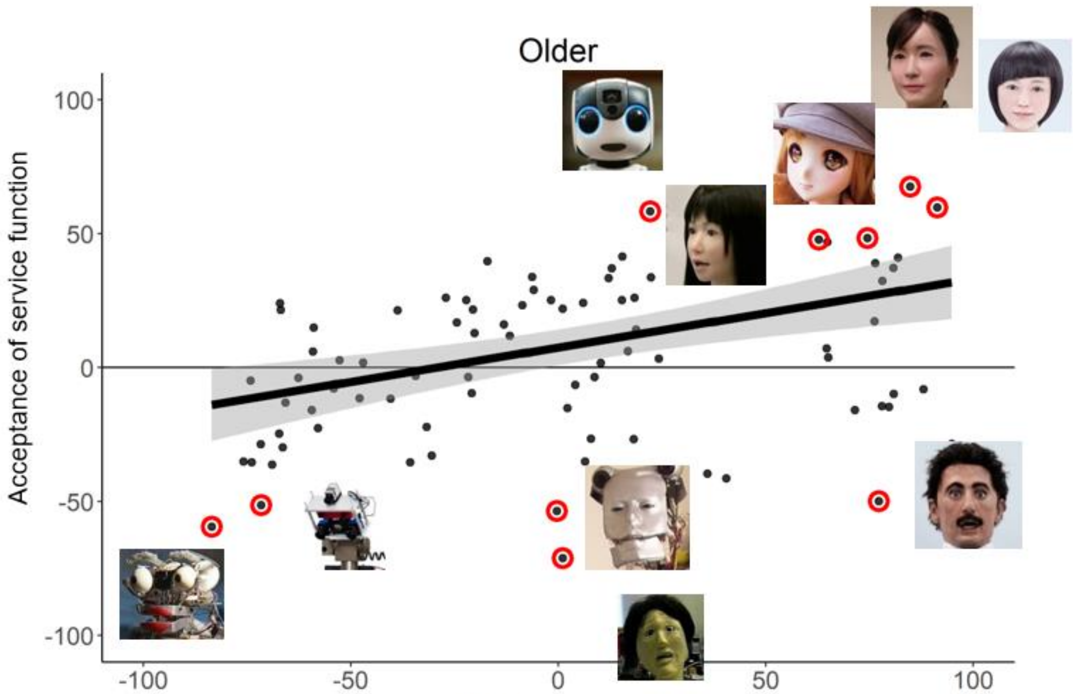

Figure 13. The top and bottom five preferable appearance for service robots based on older adults group.

Acceptance of functions, robot appearance, and perceived personalities

Trustworthiness/reliability.

Humanness.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0.00, 0.07). The rating results on trustworthiness/reliability also had low within-subject clustering (ICC = 0.06; 95% CI: 0.02, 0.11). Among all models, the third-degree model best interpreted the data (third-versus-second: $F(1, 79) = 11.748$, $p < .01$; fourth-versus-third: $F(1, 78) = 1.206$,

doi:10.6342/NTU201900877

p = .276). This model showed the UVP pattern (Figure 14, red line, R^2_{adj} = 0.141). When the humanness score reached 41.28, the trustworthiness/reliability reached the lowest score, -17.67. On the contrary, when the humanness score reached -55.692, the trustworthiness/reliability reaches the highest score, 17.48.

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The rating on the trustworthiness/reliability had low within-subject clustering (ICC = 0.2; 95% CI: 0.14, 0.27). Among all model, the third-degree model best interpreted the data (third-versus-second: $F(1, 79) = 8.662$, $p < .01$; fourth-versus-third: $F(1, 78) = 1.086$, p = .301). The model showed the UVP pattern (Figure 14, green line, R^2_{adj} = 0.160). When the humanness score reached 43.04, the rating on trustworthiness/reliability reached the lowest score, -4.30. On the contrary, when the humanness score reached -22.13, the rating on trustworthiness/reliability reached the highest turning point, 8.21.

Older adults. The result from older adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The ratings on trustworthiness/reliability also had low within-subject clustering (ICC = 0.23; 95% CI: 0.17, 0.31). Among all model, the first-degree model best interpreted the data

doi:10.6342/NTU201900877

(first-versus-no: $F(1, 81) = 8.742$, $p < .001$; second-versus-first: $F(1, 80) = 2.424$, $p = .124$; third-versus-second: $F(1, 79) = 3.829$, $p = .054$;fourth-versus-third: $F(1, 78) = 0.223$, $p = .638$). The model showed no UVP pattern but a linear relationship (Figure 14, blue line, $R^2_{adj} = 0.08$).

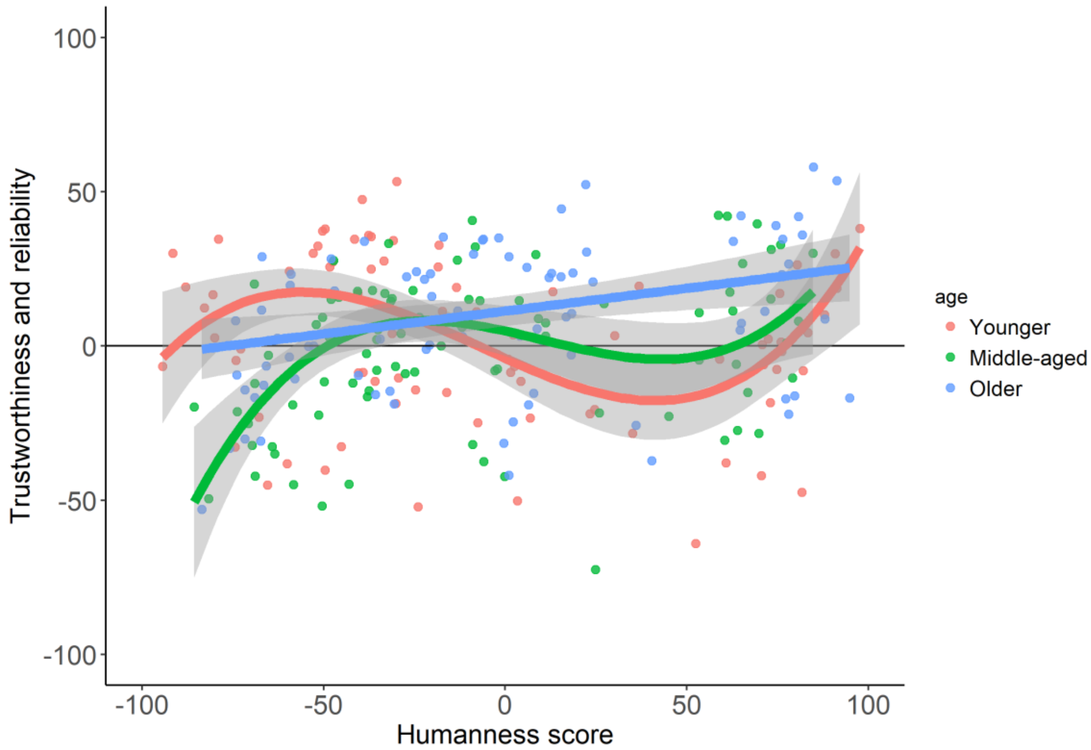

Figure 14. The relation between humanness and trustworthiness/reliability across different ages.

Companionship.

Younger adults. The result from younger adults showed that individual participants had different ratings on the trustworthiness/reliability and their ratings had low within-subject clustering (ICC = 0.06; 95% CI: 0.02, 0.11). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.2; 95% CI: 0.16, 0.27). Among all model, the first-degree model best interprets data

doi:10.6342/NTU201900877

(first-versus-no: $F(1, 81) = 275.485$, $p < .001$;second-versus-first: $F(1, 80) = 1.202$, $p = .276$;third-versus-second: $F(1, 79) = 0.056$, $p = .814$;fourth-versus-third: $F(1, 78) = 3.878$, $p = .052$), suggesting that the ratings was proportional to trustworthiness/reliability (Figure 15, red line, $R^2_{adj} = 0.765$).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the trustworthiness/reliability and their ratings had low within-subject clustering (ICC = 0.2; 95% CI: 0.14, 0.27). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.33; 95% CI: 0.27, 0.41). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 1267.285$, $p < .01$; second-versus-first: $F(1, 80) = 2.114$, $p = .15$;third-versus-second: $F(1, 79) = 2.72$, $p = .103$;fourth-versus-third: F(1, 78) = 2.303, $p = .133$), suggesting that the ratings was proportional to trustworthiness/reliability (Figure 15, green line, $R^2_{adj} = 0.936$).

Older adults. The result from older adults showed that individual participants have different ratings on the robot's trustworthiness/reliability and their ratings had low within-subject clustering (ICC = 0.23; 95% CI: 0.17, 0.31). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.37). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 776.615$, $p < .001$;second-versus-first: $F(1, 80) = 1.607$, p

doi:10.6342/NTU201900877

= .209; third-versus-second: $F(1, 79) = 0.767$, $p = 0.384$;fourth-versus-third: $F(1, 78) = 0.705$, $p = .404$), suggesting that the ratings was proportional to trustworthiness/reliability (Figure 15, blue line, $R^2_{adj} = 0.904$).

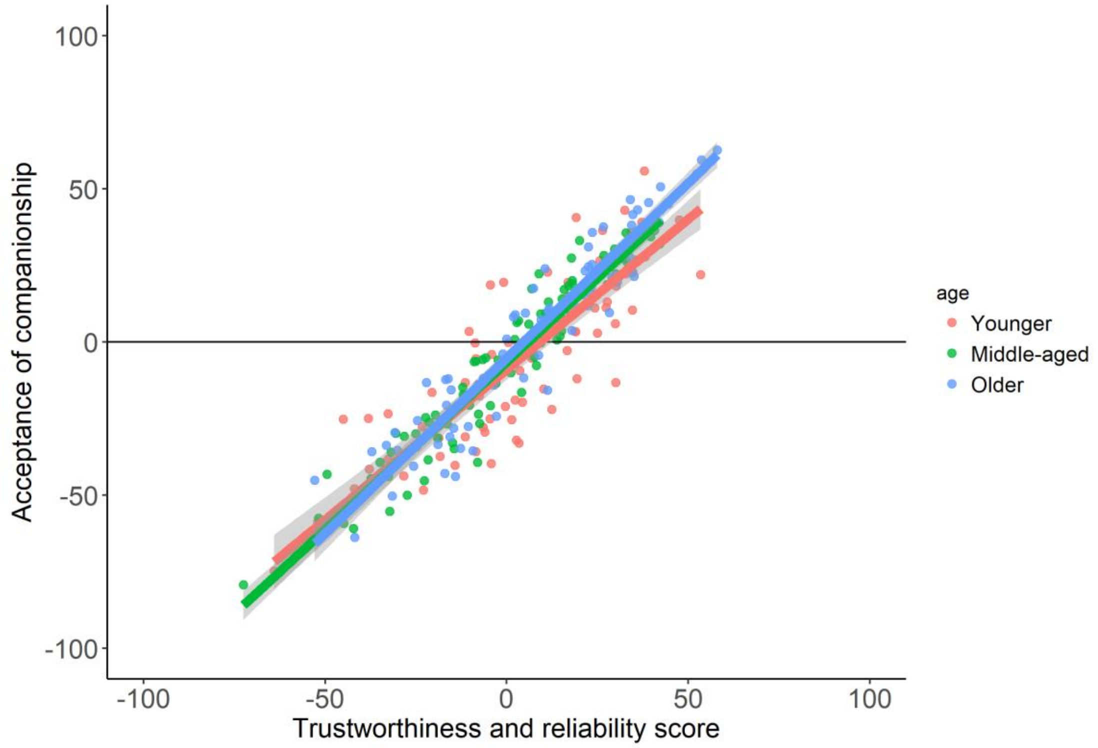

Figure 15. The relation between acceptance of companionship and trustworthiness/reliability across different age.

Service function.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's trustworthiness/reliability and their ratings had low within-subject clustering (ICC = 0.06; 95% CI: 0.02, 0.11). The rating results on the acceptance of service function also had low-within subject clustering (ICC= 0.16; 95% CI: 0.12, 0.22). Among all model, the fourth-degree model best interprets data (fourth-versus-third: $F(1, 78) = 5.905$, $p < .05$, Figure 16, red line, R^2_{adj}

doi:10.6342/NTU201900877

= 0.869), suggesting that the ratings was nearly proportional to trustworthiness/reliability (Figure 16, green line).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's trustworthiness/reliability and their ratings had low within-subject clustering (ICC = 0.2; 95% CI: 0.14, 0.27). The rating result on the acceptance of service function also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.38). Among all model, the fourth-degree model best interprets data (fourth-versus-third: $F(1, 78) = 4.362$, $p = .04$, Figure 17, red line, $R^2_{adj} = 0.913$), suggesting that the ratings was nearly proportional to trustworthiness/reliability (Figure 17, red and green line).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's trustworthiness/reliability and their ratings had low within-subject clustering (ICC = 0.23; 95% CI: 0.17, 0.31). The rating result on the acceptance of service function also had low-within subject clustering (ICC= 0.28; 95% CI: 0.23, 0.36). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 675.791$, $p < .001$;second-versus-first: $F(1, 80) = 0.165$, $p = .686$; third-versus-second: $F(1, 79) = 0.047$, $p = .829$;fourth-versus-third: $F(1, 78) = 0.406$, $p = .526$), suggesting that the ratings was proportional to

doi:10.6342/NTU201900877

trustworthiness/reliability (Figure 18, blue line, $R^2_{adj} = 0.894$).

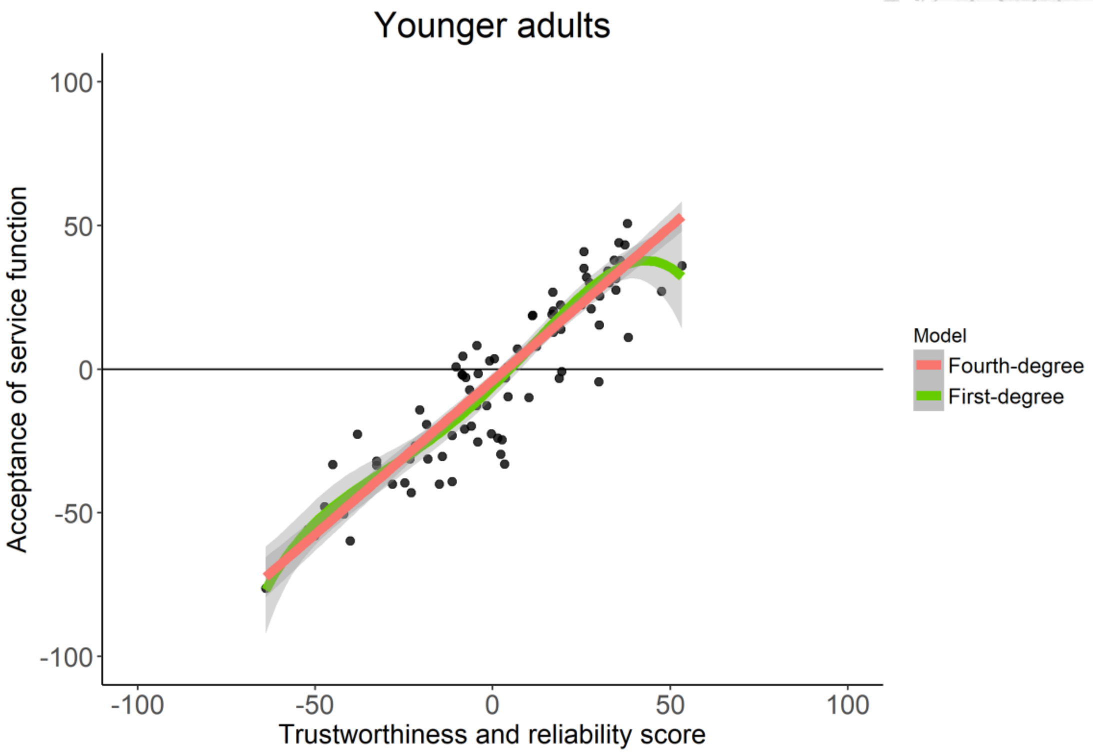

Figure 16. The relation between acceptance of service function and trustworthiness/reliability of younger adults.

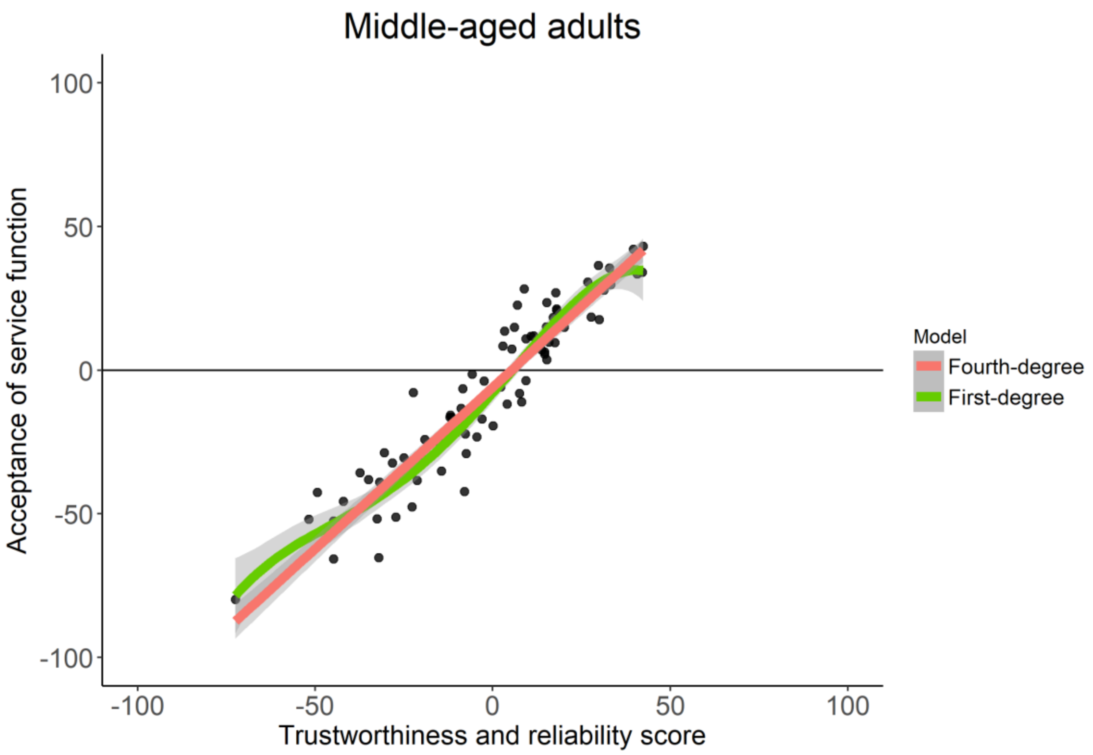

Figure 17. The relation between acceptance of service function and trustworthiness/reliability of middle-aged adults.

doi:10.6342/NTU201900877

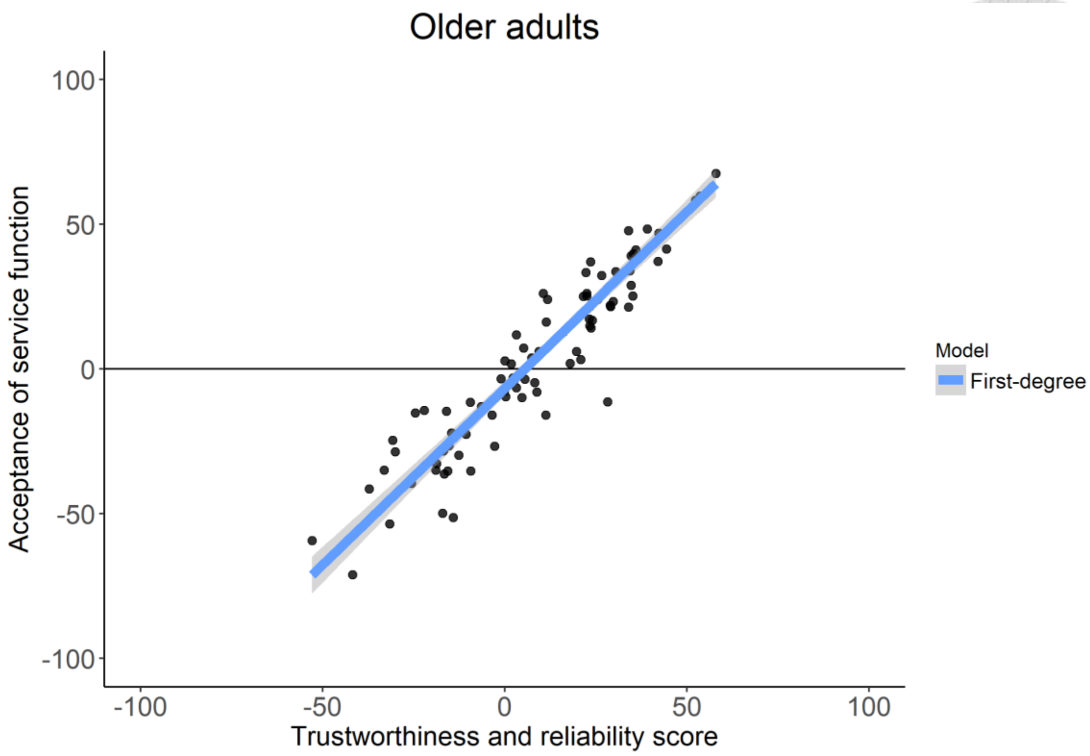

Figure 18. The relation between acceptance of service function and trustworthiness/reliability of older adults.

Friendliness/approachability.

Humanness.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0.00, 0.07). The ratings on the robot's friendliness/approachability also had low within-subject clustering (ICC = 0.07; 95% CI: 0.03, 0.12). Among all models, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 19.161$, $p < .001$;fourth-versus-third: $F(1, 78) = 0.036$, $p = .851$). This model showed the UVP pattern (Figure 19, red line, $R^2_{adj} = 0.179$). When the humanness score reached 43.63, the friendliness/approachability reached

doi:10.6342/NTU201900877

the lowest score, -20.71. On the contrary, when the humanness score reached -48.29, the ratings on friendliness/approachability reaches the highest score, 22.05.

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The ratings on robot's friendliness/approachability also had low within-subject clustering (ICC = 0.17; 95% CI: 0.12, 0.24). Among all model, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 9.765$, $p < .01$;fourth-versus-third: $F(1, 78) = 1.199$, $p = .277$). The model showed the UVP pattern (Figure 19, green line, $R^2_{adj} = 0.174$). When the humanness score reached 42.692, the rating on friendliness/approachability reached the lowest score, -6.497. On the contrary, when the humanness score reached -24.00, the rating on friendliness/approachability reached the highest turning point, 9.55.

Older adults. The result from older adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The ratings on the robot's friendliness/approachability also had low within-subject clustering (ICC = 0.18; 95% CI: 0.12, 0.25). Among all model, the second-degree model best interprets data (second-versus-first: $F(1, 80) = 4.285$, $p = .042$; third-versus-second: $F(1, 79) =$

doi:10.6342/NTU201900877

2.964, p = .089; fourth-versus-third: $F(1, 78) = 1.237$, p = .270). The model showed no UVP pattern (Figure 19, blue line, $R^2_{adj} = 0.126$).

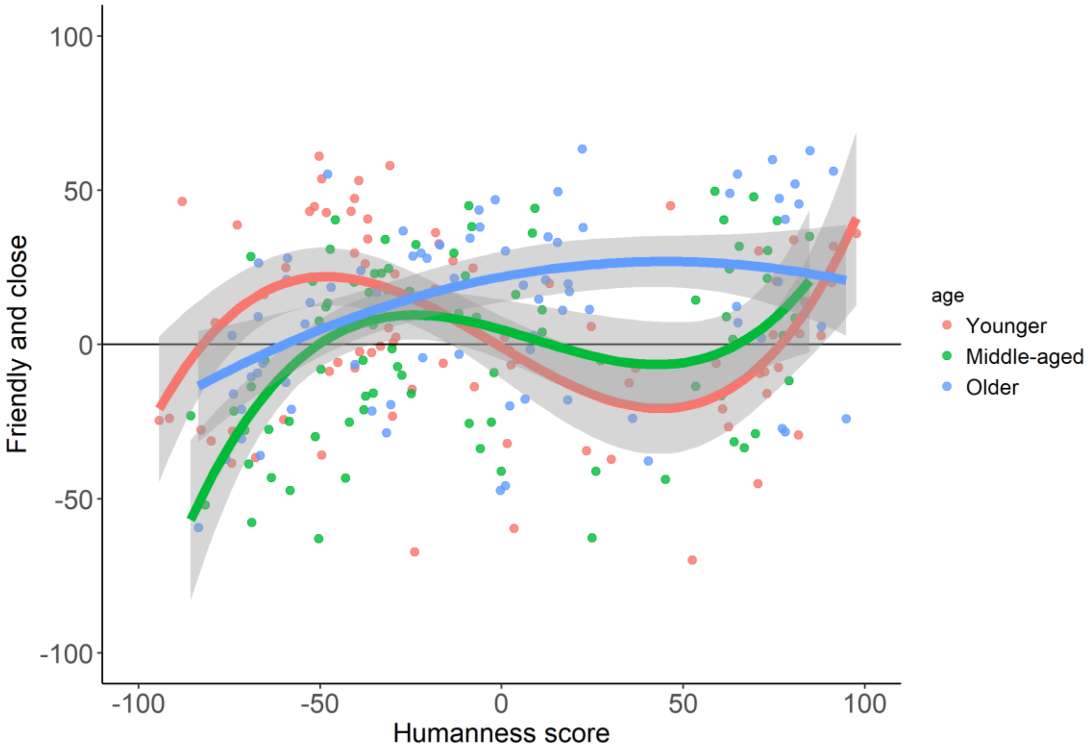

Figure 19. The relation between humanness and friendliness/approachability across different ages.

Companionship.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's friendliness/approachability and their ratings had low within-subject clustering (ICC = 0.07; 95% CI: 0.03, 0.12). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.2; 95% CI: 0.16, 0.27). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 371.851$, $p < .001$; second-versus-first: $F(1, 80) = 0.485$, p = .488; third-versus-second: $F(1, 79) = 0.576$, p = .45;fourth-versus-

doi:10.6342/NTU201900877

third: $F(1, 78) = 0.123$, $p = .727$), suggesting that the ratings was proportional to friendliness/approachability (Figure 20, red line, $R^2_{adj} = 0.822$).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's friendliness/approachability and their ratings had low within-subject clustering (ICC = 0.17; 95% CI: 0.12, 0.24). The rating results on the robots' companion score also had low-within subject clustering (ICC= 0.33; 95% CI: 0.27, 0.41). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 1371.902$, $p < .01$; second-versus-first: $F(1, 80) = 1.07$, $p = .304$;third-versus-second: $F(1, 79) = 1.406$, $p = .239$;fourth-versus-third: $F(1, 78) = 0.027$, $p = .87$) suggesting that the ratings was proportional to friendliness/approachability (Figure 20, green line, $R^2_{adj} = 0.944$).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's friendliness/approachability and their ratings had low within-subject clustering (ICC = 0.18; 95% CI: 0.12, 0.25). The rating results on the acceptance of companionship also have low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.37). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 871.167$, $p < .001$; second-versus-first: $F(1, 80) = 0.131$, $p = .718$; third-versus-second: $F(1, 79) = 2.3$, $p = .133$;fourth-versus-third: $F(1, 78) = 0.676$, $p = .413$), suggesting that the ratings was proportional to

doi:10.6342/NTU201900877

friendliness/approachability (Figure 20, blue line, $R^2_{adj} = 0.914$).

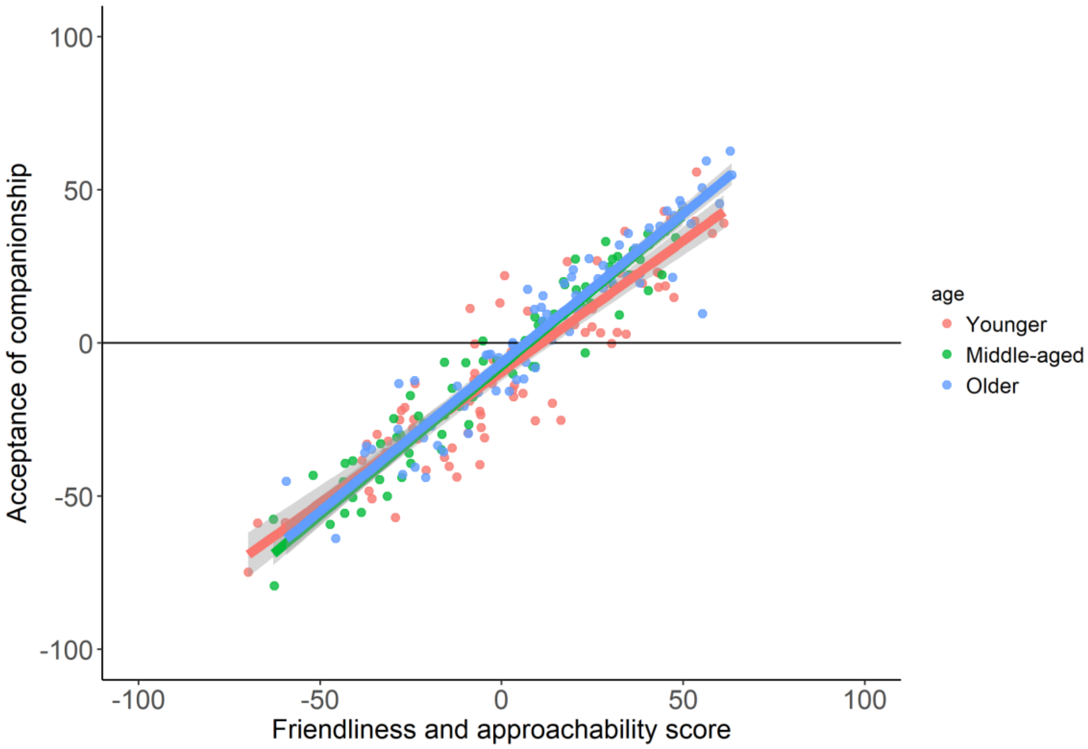

Figure 20. The relation between acceptance of companionship and perceived friendliness/approachability of the robot across different ages.

Service function.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's friendliness/approachability and their ratings had low within-subject clustering (ICC = 0.07; 95% CI: 0.03, 0.12). The rating results on the robots' service score also had low-within subject clustering (ICC= 0.16; 95% CI: 0.12, 0.22). Among all models, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 152.849$, $p < .001$; second-versus-first: $F(1, 80) = 2.066$, $p = .155$; third-versus-second: $F(1, 79) = 0.216$, $p = .644$;fourth-versus-third: $F(1, 78) = 0.074$, $p = .786$), suggesting that the rating was proportional to

doi:10.6342/NTU201900877

friendliness/approachability (Figure 21, red line, R^2_{adj} = 0.651).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's friendliness/approachability and their ratings had low within-subject clustering (ICC = 0.17; 95% CI: 0.12, 0.24). The rating results on the robots' service score also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.38). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 789.974$, $p < .001$; second-versus-first: $F(1, 80) = 0.105$, $p = .747$; third-versus-second: $F(1, 79) = 2.428$, $p = .123$;fourth-versus-third: $F(1, 78) = 0.007$, $p = .935$), suggesting that the rating was proportional to friendliness/approachability (Figure 21, green line, R^2_{adj} = 0.906).

Older adults. The result from older adults shows that individual participants had different ratings on the robot's friendliness/approachability and their ratings had low within-subject clustering (ICC = 0.18; 95% CI: 0.12, 0.25). The rating results on the robots' service score also had low-within subject clustering (ICC= 0.28; 95% CI: 0.23, 0.36). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 609.409$, $p < .01$; second-versus-first: $F(1, 80) = 0.248$, p = .62; third-versus-second: $F(1, 79) = 0.37$, $p = .545$;fourth-versus-third: $F(1, 78) = 0.439$, $p = .51$), suggesting that the ratings was proportional to friendliness/approachability

doi:10.6342/NTU201900877

(Figure 21, blue line, R^2_{adj} = 0.884).

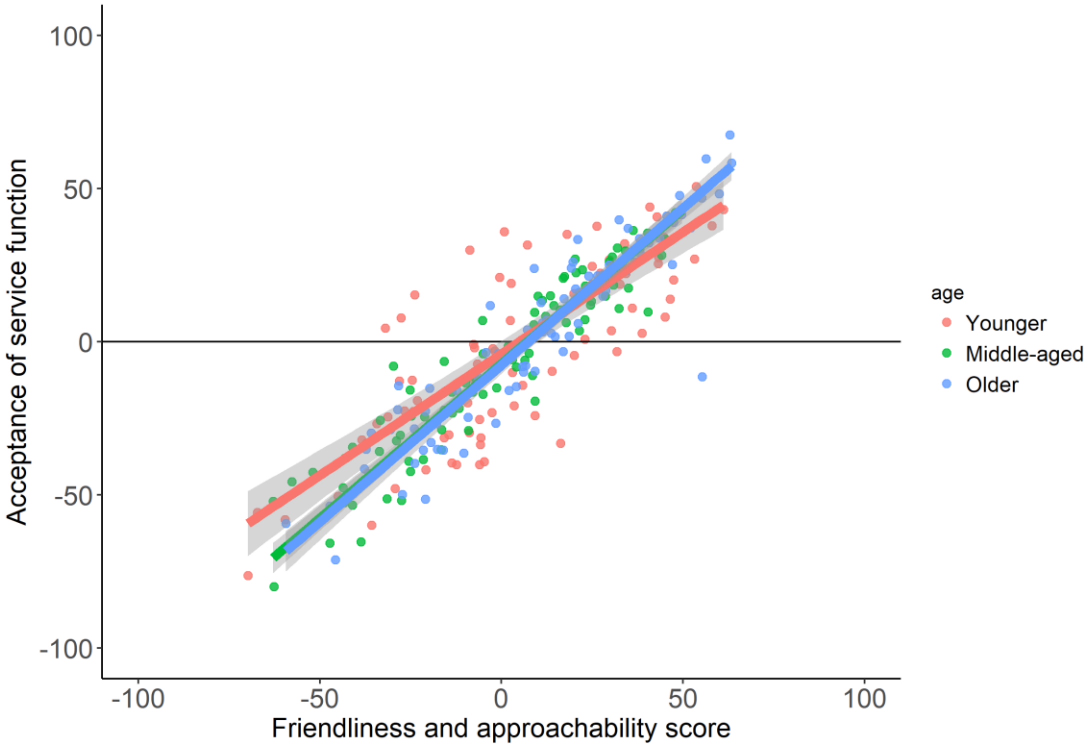

Figure 21. The relation between acceptance of service function and friendliness/approachability across different age.

Authoritativeness.

Humanness.

*Younger adults***.** The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0.00, 0.07). The ratings on the robot's authoritativeness also had low within-subject clustering (ICC = 0.13; 95% CI: 0.08, 0.19). Among all models, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 11.362$, $p < .001$; second-versus-first: $F(1, 80) = 0.013$, $p = .910$;third-versus-second: $F(1, 79) = 0.949$, $p = .332$;fourth-versus-third: $F(1, 78) = 0.660$, p

doi:10.6342/NTU201900877

= .419). This model showed no UVP pattern but a linear relationship (Figure 22, red line, R^2_{adj} = 0.114).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The ratings on the robot's authoritativeness also had low within-subject clustering (ICC = 0.23; 95% CI: 0.17, 0.3). Among all models, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 32.705$, $p < .001$; second-versus-first: $F(1, 80) = 2.180$, $p = .144$;third-versus-second: $F(1, 79) = 2.403$, $p = .125$;fourth-versus-third: $F(1, 78) = 0.086$, $p = .770$). The model showed no UVP pattern but a linear relationship (Figure 22, green line, R^2_{adj} = 0.274).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The ratings on the robot's authoritativeness also had low within-subject clustering (ICC = 0.24; 95% CI: 0.18, 0.32). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 5.926$, $p < .017$; second-versus-first: $F(1, 80) = 0.044$, $p = .834$; third-versus-second: $F(1, 79) = 1.319$, $p = .254$; fourth-versus-third: $F(1, 78) = 0.543$, $p = .463$). The model showed no UVP pattern but a linear relationship (Figure 22, blue

doi:10.6342/NTU201900877

line, R^2_{adj} = 0.058).

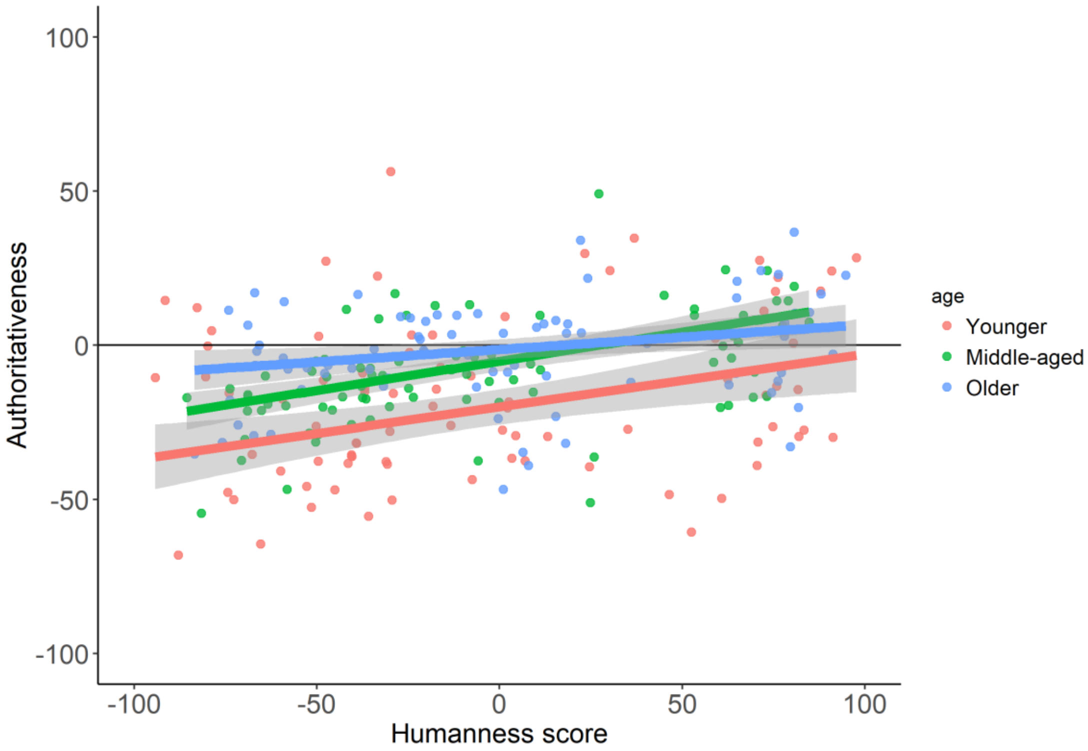

Figure 22. The relation between humanness and authoritative across different ages.

Companionship.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's authoritativeness and their ratings had low within-subject clustering (ICC = 0.13; 95% CI: 0.08, 0.19). The rating results on the acceptance of companionship also had low-within subject clustering (ICC= 0.2; 95% CI: 0.16, 0.27). Among all models, the horizontal line best interprets the data (first-versus-no: $F(1, 81) = 0.026$, $p = .872$; second-versus-first: $F(1, 80) = 0.048$, $p = .828$; third-versus-second: $F(1, 79) = .47$, $p = 0.495$;fourth-versus-third: $F(1, 78) = 0.698$, $p = .406$), suggesting that that the ratings maintained equal no matter what

doi:10.6342/NTU201900877

scores of authoritativeness were (Figure 23, red line).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's authoritativeness and their ratings had low within-subject clustering (ICC = 0.23; 95% CI: 0.17, 0.3). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.33; 95% CI: 0.27, 0.41). Among all models, the second-degree model performs best (second-versus-first: $F(1, 80) = 7.833$, $p < .01$; third-versus-second: $F(1, 79) = 0.012$, $p = .912$;fourth-versus-third: $F(1, 78) = 1.074$, $p = .303$), suggesting that when ratings of authoritativeness scores reached middle-high range, 19.18, the acceptance of service function reached the lowest score, 7.95. (Figure 23, green line, R^2_{adj} =0.264).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's authoritativeness and their ratings had low within-subject clustering (ICC = 0.24; 95% CI: 0.18, 0.32). The rating results of the acceptance of companionship also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.37). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 26.371$, $p < .001$; second-versus-first: $F(1, 80) = 0.674$, $p = .414$; third-versus-second: $F(1, 79) = 2.408$, $p = .125$;fourth-versus-third: $F(1, 78) = 0.001$, $p = .976$), suggesting that that the ratings were proportional to scores of

doi:10.6342/NTU201900877

authoritativeness (Figure 23, blue line, R^2_{adj} =0.236).

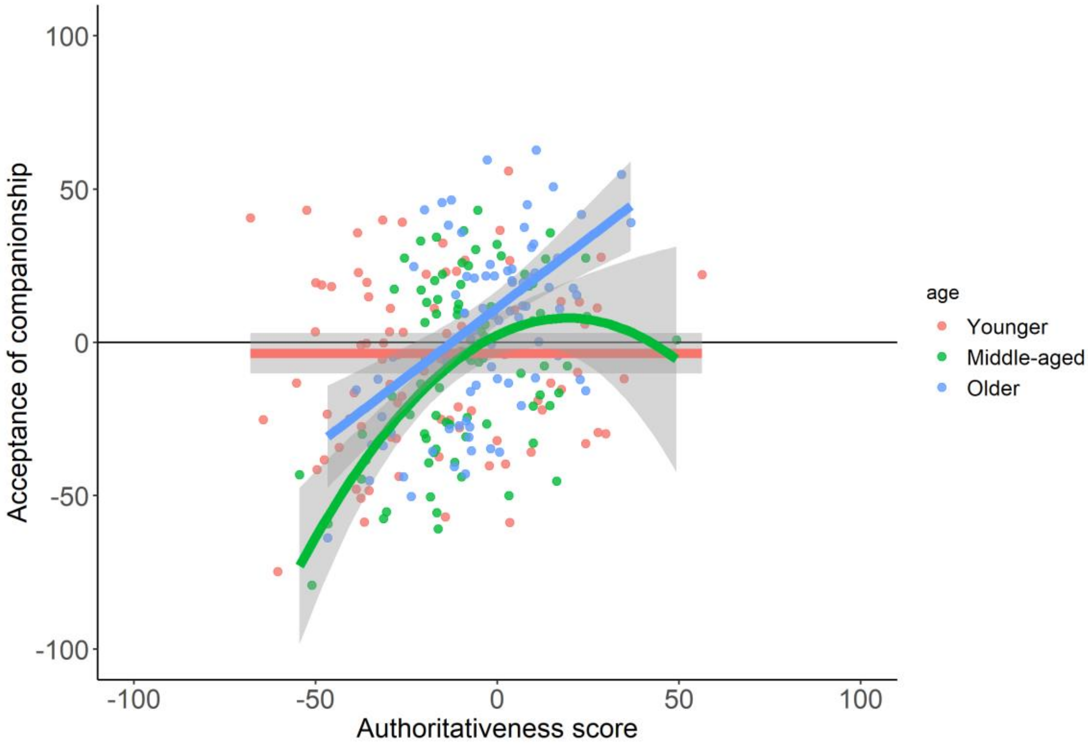

Figure 23. The relation between acceptance of companionship and authoritativeness across different age groups.

Service function.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's authoritativeness and their ratings had low within-subject clustering (ICC = 0.13; 95% CI: 0.08, 0.19). The rating result on the acceptance of service function also had low-within subject clustering (ICC= 0.16; 95% CI: 0.12, 0.22). Among all model, the first-degree model best interprets data (first-versus-no: F(1, 81) = 6.094, $p < .05$; second-versus-first: F(1, 80) = 2.172, $p = .145$; third-versus-second: F(1, 79) = 1.391, $p = .242$;fourth-versus-third: F(1, 78) = 0.752, $p = .389$), suggesting that that the ratings were proportional to scores of

doi:10.6342/NTU201900877

authoritativeness (Figure 24, red line, $R^2_{adj} = 0.057$).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's authoritativeness and their ratings had low within-subject clustering (ICC = 0.23; 95% CI: 0.17, 0.3). The rating result on the robots' service score also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.38). Among all model, the second-degree model best interprets data (second-versus-first: $F(1, 80) = 8.135$, $p < .01$;third-versus-second: $F(1, 79) = 0.017$, $p = .896$;fourth-versus-third: $F(1, 78) = 1.473$, $p = .229$), suggesting that when ratings of authoritativeness scores reached middle-high range, 21.65, the acceptance of service function reached the lowest score, 10.86 (Figure 24, green line, $R^2_{adj} = 0.296$).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's authoritativeness and their ratings had low within-subject clustering (ICC = 0.24; 95% CI: 0.18, 0.32). The rating result on the robots' service score also had low-within subject clustering (ICC= 0.28; 95% CI: 0.23, 0.36). Among all models, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 32.222$, $p < 0.001$; second-versus-first: $F(1, 80) = 2.214$, $p = .141$; third-versus-second: $F(1, 79) = 3.616$, $p = .061$;fourth-versus-third: $F(1, 78) = 0.064$, $p = .801$), suggesting that that the ratings were proportional to scores of authoritativeness

doi:10.6342/NTU201900877

(Figure 24, blue line, $R^2_{adj} = 0.269$).

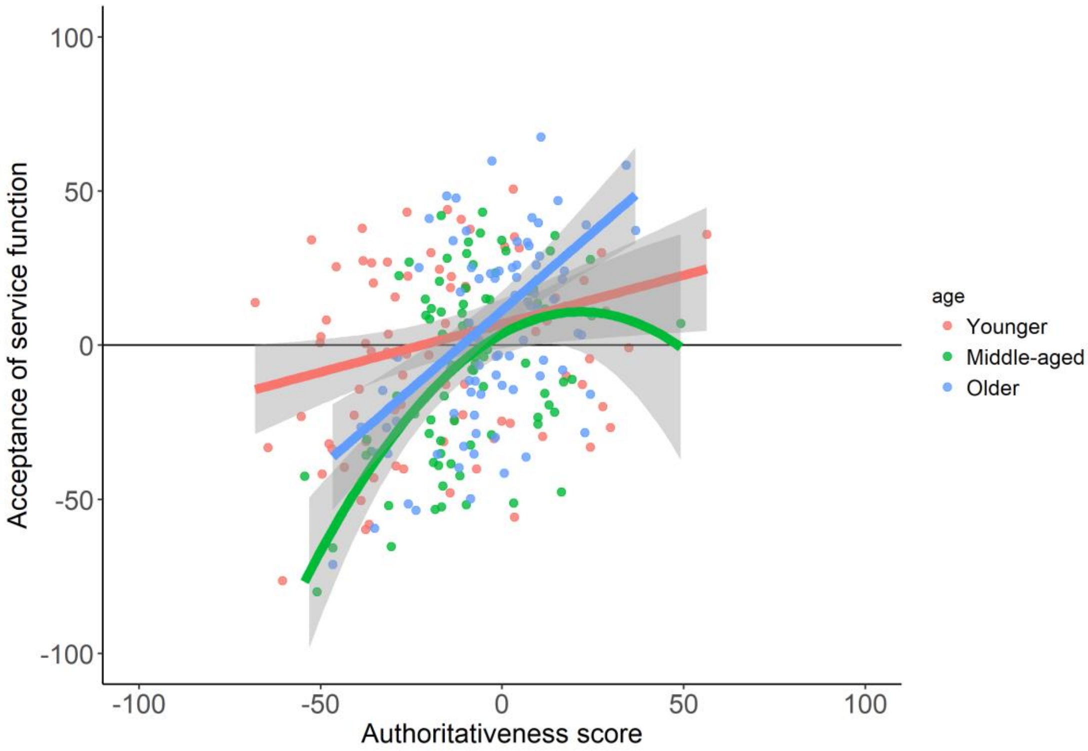

Figure 24. The relation between acceptance of service function and authoritative across different age.

Adorability.

Humanness.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.03; 95% CI: 0.00, 0.07). The ratings on the robot's adorability also had low within-subject clustering (ICC = 0.04; 95% CI: 0.01, 0.08). Among all models, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 8.426$, $p < .01$;fourth-versus-third: $F(1, 78) = 2.069$, $p = .154$). This model showed the UVP pattern (Figure 25, red line, $R^2_{adj} = 0.121$). When the humanness

doi:10.6342/NTU201900877

score reached 44.68, the adorability reached the lowest score, -54.15. On the contrary, when the humanness score reached -56.25, the adorability reaches the highest score, 1.79.

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.12; 95% CI: 0.07, 0.18). The ratings on the robot's adorability also had low within-subject clustering (ICC = 0.16; 95% CI: 0.11, 0.22). Among all model, the third-degree model best interprets data (third-versus-second: $F(1, 79) = 9.582$, $p < .01$;fourth-versus-third: F(1, 78) = 0.943, $p = .334$). The model showed the UVP pattern (Figure 25, green line, $R^2_{adj} = 0.163$). When the humanness score reached 44.68, the adorability reached the lowest score, -54.15. On the contrary, when the humanness score reached -56.26, the adorability reached the highest turning point, 1.79.

Older adults. The result from older adults showed that individual participants had different ratings on the robot's humanness and their ratings had low within-subject clustering (ICC = 0.09; 95% CI: 0.06, 0.15). The ratings on the robot's adorability also had low within-subject clustering (ICC = 0.13; 95% CI: 0.09, 0.2). Among all model, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 9.786$, $p < .001$;second-versus-first: $F(1, 80) = 1.454$, $p = .231$;third-versus-second:

doi:10.6342/NTU201900877

$F(1, 79) = 3.771$, $p = .055$;fourth-versus-third: $F(1, 78) = 1.379$, $p = .243$). The model showed no UVP pattern but a linear relationship (Figure 25, blue line, $R^2_{adj} = 0.09$).

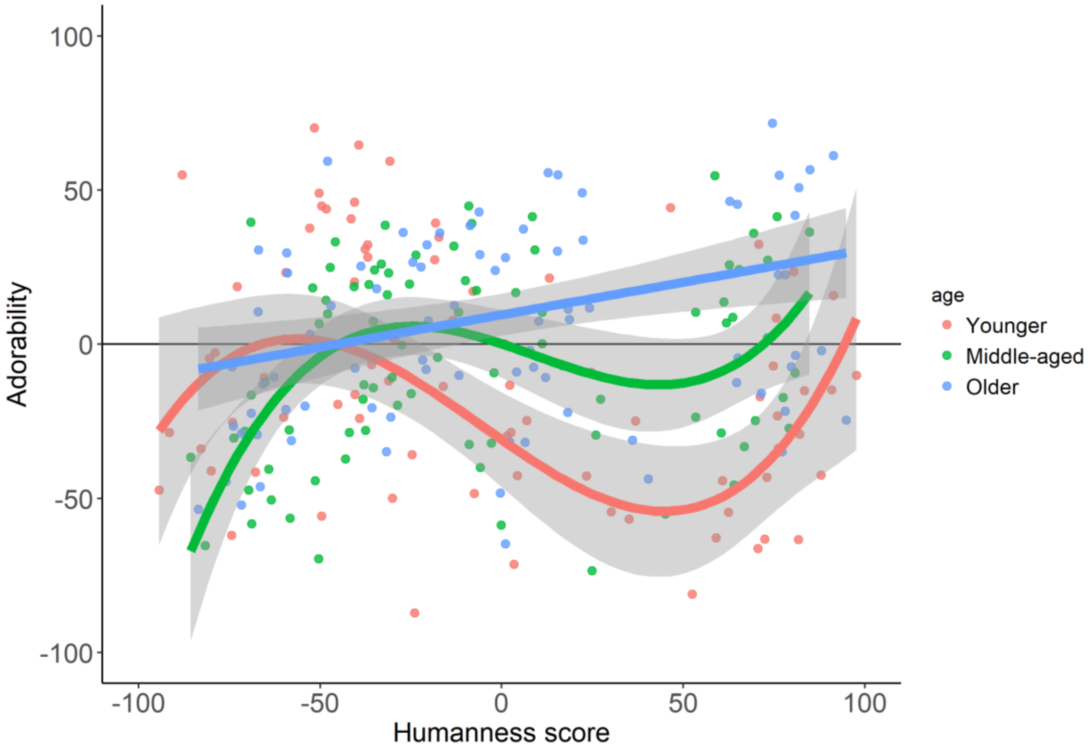

Figure 25. The relation between humanness and adorability across different ages.

Companionship.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's adorability and their ratings had low within-subject clustering (ICC = 0.04; 95% CI: 0.01, 0.08). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.2; 95% CI: 0.16, 0.27). Among all models, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 553.231$, $p < .01$; second-versus-first: $F(1, 80) = 1.672$, $p = .20$; third-versus-second: $F(1, 79) = 0.122$, $p = .727$;fourth-versus-third: $F(1, 78) = 0.502$, $p = .481$), suggesting that that the ratings were proportional to scores of adorability

doi:10.6342/NTU201900877

(Figure 26, red line, R^2_{adj} = 0.872).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's adorability and their ratings had low within-subject clustering (ICC = 0.16; 95% CI: 0.11, 0.22). The rating result on the acceptance of companionship also had low-within subject clustering (ICC= 0.33; 95% CI: 0.27, 0.41). Among all model, the first-degree model best interprets data (first-versus-no: F(1, 81) = 904.188, $p < .001$; second-versus-first: F(1, 80) = 0.264, $p = .609$; third-versus-second: F(1, 79) = 0.264, $p = .609$;fourth-versus-third: F(1, 78) = 0.048, $p = .827$), suggesting that that the ratings were proportional to scores of adorability (Figure 26, green line, R^2_{adj} = 0.919).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's adorability and their ratings had low within-subject clustering (ICC = 0.13; 95% CI: 0.09, 0.2). The rating results of the acceptance of companionship also have low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.37). Among all models, the second-degree model best interprets data (second-versus-first: F(1, 80) = 4.206, $p < .05$;third-versus-second: F(1, 79) = 0.421, $p = .518$; fourth-versus-third: F(1, 78) = 0.161, $p = .689$), suggesting that that the ratings were

doi:10.6342/NTU201900877

proportional to scores of adorability (Figure 26, blue line, $R^2_{adj} = 0.884$).

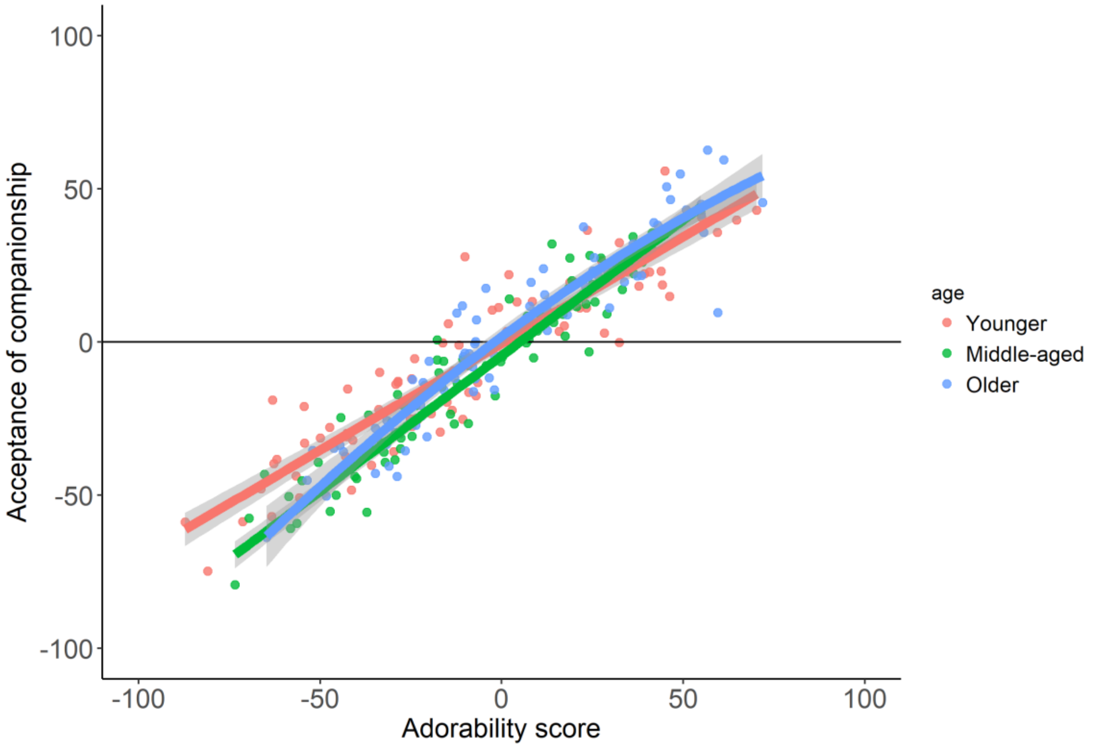

Figure 26. The relation between acceptance of companionship and adorability across different age.

Service function.

Younger adults. The result from younger adults showed that individual participants had different ratings on the robot's adorability and their ratings had low within-subject clustering (ICC = 0.04; 95% CI: 0.01, 0.08). The rating result on the acceptance of service function also had low-within subject clustering (ICC= 0.16; 95% CI: 0.12, 0.22). Among all models, the second-degree model best interprets data (second-versus-first: $F(1, 80) = 11.759$, $p < .01$; third-versus-second: $F(1, 79) = 0.965$, $p = .329$;fourth-versus-third: $F(1, 78) = 0.508$, $p = .478$). Also, it suggests that there is positive correlation between ratings of adorability scores and acceptance of service function. (Figure 27, red line, R^2_{adj} =0.757).

Middle-aged adults. The result from middle-aged adults showed that individual participants had different ratings on the robot's adorability and their ratings had low within-subject clustering (ICC = 0.16; 95% CI: 0.11, 0.22). The rating results on the acceptance of service function also had low-within subject clustering (ICC= 0.3; 95% CI: 0.24, 0.38). Among all models, the first-degree model best interprets data (first-versus-no: $F(1, 81) = 674.197$, $p < .001$; second-versus-first: $F(1, 80) = 0.062$, $p = .804$;third-versus-second: $F(1, 79) = 0.815$, $p = .369$;fourth-versus-third: F(1, 78) = 0.007, $p = .932$), suggesting that that the ratings were proportional to scores of adorability (Figure 27, green line, $R^2_{adj} = 0.894$).

Older adults. The result from older adults showed that individual participants had different ratings on the robot's adorability and their ratings had low within-subject clustering (ICC = 0.13; 95% CI: 0.09, 0.2). The rating results on the acceptance of service function also had low-within subject clustering (ICC= 0.28; 95% CI: 0.23, 0.36). Among all models, the second-degree model best interprets data (second-versus-first: F(1, 80) = 4.172, $p < .05$; third-versus-second: F(1, 79) = 0.027, $p = .869$; fourth-versus-third: F(1, 78) = 0.31, $p = .579$). Also, it suggests that there is positive correlation between ratings of adorability scores and acceptance of service function

doi:10.6342/NTU201900877

(Figure 27, blue line, $R^2_{adj} = 0.836$).

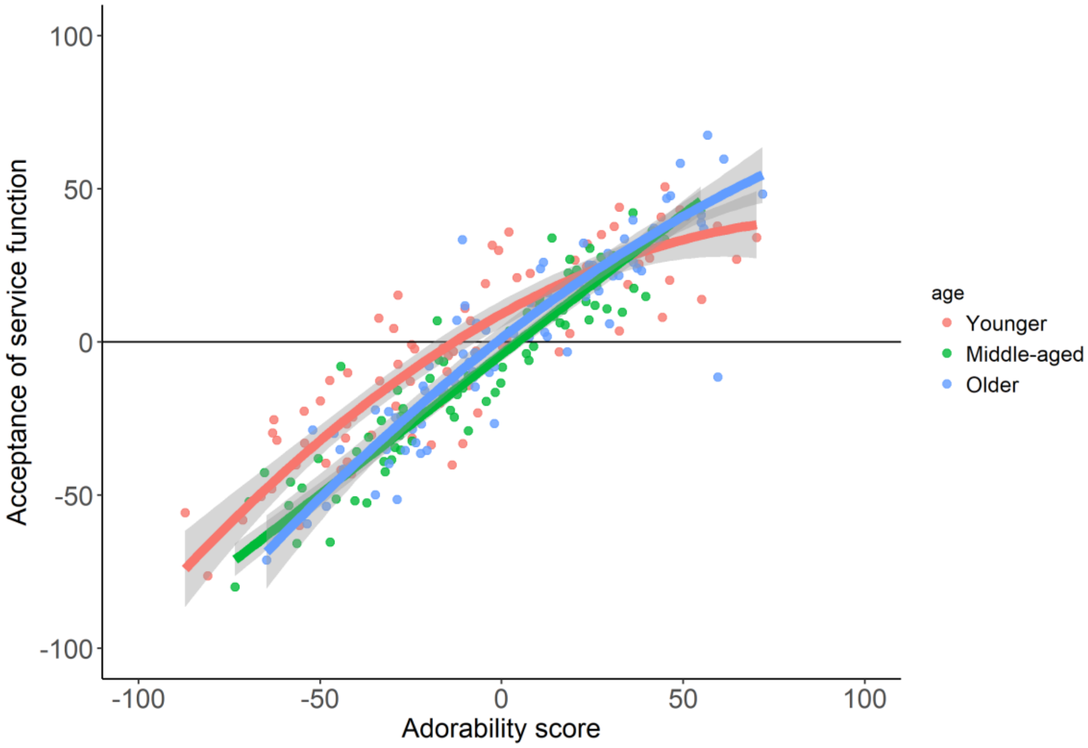

Figure 27. The relation between acceptance of service function and adorability across different age.

doi:10.6342/NTU201900877

General Discussion

In our study, we used questionnaires to answer three main questions: (1) whether the uncanny valley phenomenon varied across different age groups, (2) whether the acceptance of companion or service functions would change based on robot appearance, and what would be an ideal design of robot appearance according to the function a robot would serve, (3) whether robot-perceived personalities of robots could have any relation with the acceptance of companionship and service functions. According to our results, the uncanny valley phenomenon did vary across different age groups. Also, users in different age groups would change their acceptance of robots' based on robots' appearance. Besides, users are able to perceive robots' personality because of robots' appearance.

The uncanny valley phenomenon across different age groups

The major finding of the present study is that the UVP could not be observed across all age groups. The UVP only existed in younger and middle-aged adults groups, but there was no UVP for older adults. Older adults showed a positive linear correlation between likeability and humanness scores. Although both younger and middle-aged adults showed the UVP, younger adults showed preference for non-humanlike robots while middle-aged adults showed preference for humanlike robots (Figure 3-5). Also, our results from younger adults was in conformity with results

doi:10.6342/NTU201900877

from Mathur and Reichling (2016) that participants mostly belonged to younger adults group in our study, indicating the three-degree model could represent the relation between humanness scores and likeability. Thus, we should take users' age into consideration as we discuss the UVP in the future, for not all users would display this tendency.

Our results demonstrated that robot appearance had strong influence on users' likeability. Hence, when designing robots, we should take preference for appearance of our target audience into consideration. Designers should be aware of what age their clients might be, so they could properly design robots' appearances for users. On the other hand, for those who may want to buy robots for their parents or children, choosing robots with different appearances based on the age of potential users would be critical to increase the user acceptance of the robot.

Age could be an important factor of the UVP. Lewkowicz and Ghazanfar (2012) found that infants showed preference for human faces after 12 months. The authors suggested that the capability to perceive details of human faces increases rapidly between 6-12 months after birth. Tung (2011) found that children aged from 8-13 started to show different UVP patterns based on their genders: Boys reported higher social and physical attractiveness to non-humanlike robots than girls did, while girls reported higher social and physical attractiveness to humanlike robots than boys did.

doi:10.6342/NTU201900877

However, no studies except the current one, as far as we know, has extended the UVP to the higher end of the entire life span.

Our result showed that age not only affected the starting point of the UVP, but also affected whether UVP existed: the UVP vanished in older adults group. *Socioemotional selectivity theory* indicates that older adults would put much emphasis on emotional regulation than both younger and middle-aged adults (Carstensen, 1995), so we could predict that older adults should prefer robots that could provide emotional support for them. Most pictures in our study belong to either the category of humanlike or mechanical robots. In other words, older adults would assume that those non-humanlike or mechanical robots were unable to provide emotional support, which resulted in lower likeability.

Two hypotheses for the UVP may explain our results. First, the *expectation violation hypothesis,* which states that people would rapidly label one humanlike robot into "human" model and use this "human" model to check the robot. When the robot fails to satisfy some properties of the model, people can discriminate the difference between humanlike robots and real human (MacDorman & Ishiguro, 2006). However, these subtle differences would induce the feeling of unpredictability, which would make people feel weird (Saygin, Chaminade, Ishiguro, Driver, & Frith, 2011). As Cheng, Shyi, and Cheng (2016) suggested that older adults' strategy to

doi:10.6342/NTU201900877

discriminate facial stimuli would be different from younger adults'; older adults tend to process face stimuli in the holistic way while younger adults would process face stimuli in both holistic and analytic ways. It is possible that older adults have lower sensitivity of these subtle differences when viewing robots pictures, and thus resulted in no intensive negative emotions. Besides, MacDorman and Entezari (2015) indicated that individual differences such as the ability to discriminate the living and non-living things and the anxiety traits are both positively correlated with their feelings of eerie toward robots, and both the two factors are negatively correlated with users' age. That is, older adults might ignore the conflictory cues from nearly humanlike robots and have lower anxiety to those robots. Thus, participants didn't show lower preference on those nearly humanlike robots.

Secondly, the *Mortality salience hypothesis* suggests that people have innate fear of death. Some robots might be a reminder of death, so people when see those robots would experience negative emotion toward them (MacDorman & Ishiguro, 2006). Indeed, older adults tended to show lower disgusted feelings toward diseases-and-dead-related pictures (Curtis, Aunger, & Rabie, 2004), which could be one possible explanation of why older adults did not show the UVP: the robots induced lower negative feelings to them.

doi:10.6342/NTU201900877

Acceptance of functions and appearance

Our results showed that the correlation between the likeability toward robots and their degree of humanness was similar to the correlation between acceptance of both companionship and service functions and robot humanness. While both younger and middle-aged adults showed the UVP, older adults did not. Similarly, younger adults would accept assistance from both humanlike and non-humanlike robots, while older adults prefer humanlike robots. In addition, although middle-aged adults showed the UVP (third-degree model), they still preferred humanlike robots as older adults did (Figure 7-9,11-12). These results indicated that we could design specific appearance of robots according to users' ages to increase their acceptance of robot's assistance. We also demonstrated the best and worst five appearances as the recommendations for designers (Figure 7-9,11-12). The results from pictures clearly showed that younger adults preferred to be served by robots with less humanlike appearance, while middle-aged and older adults chose more human-like robots. These figures can be used as references for future robotic designs.

Again, our results could be explained by the *socioemotional selectivity theory* (Carstensen, 1995). Younger and middle-aged adults had lower need for emotional regulation than older adults (Carstensen, 1995). Thus, not all users showed higher preference on humanlike robots. However, there is still a difference worth discussing.

doi:10.6342/NTU201900877

Based on previous studies, older adults would choose non-humanlike robots to help them doing housework and moving the objects and yet they preferred to be taken care of and accompanied by humanlike robots (Smarr et al., 2014; Smarr et al., 2012). In addition, Chu et al. (2019) indicated that older adults preferred service-oriented robots. At the first glance, our results demonstrated that older adults' acceptance of companionship were similar to service functions (Figure 28). One possibility might be that the functions in the present study were selected according to Chu et al. (2019) that identified ideal functions of robots. In other words, we only selected most wanted functions for older adults. Nonetheless, we can't compare them because two function models came from different data, and there is no statistic method to examine these difference.

Acceptance of functions, robot appearance, and personalities

Our data showed that even a static image of the robot was enough for users to make a preliminary judgment on the robot's personality, which was congruent with results from Lee et al. (2006). They suggested that people could recognize pet robots' personalities by observing robots' verbal (voice, melody) and non-verbal signals (led lights, moving angle, moving speed, etc.). Indeed, people can rapidly label personalities on those who they are not acquainted with, and these rapidly made judgements on extraversion and conscientiousness are highly congruent with those

doi:10.6342/NTU201900877

being carefully evaluated (Albright et al., 1988).

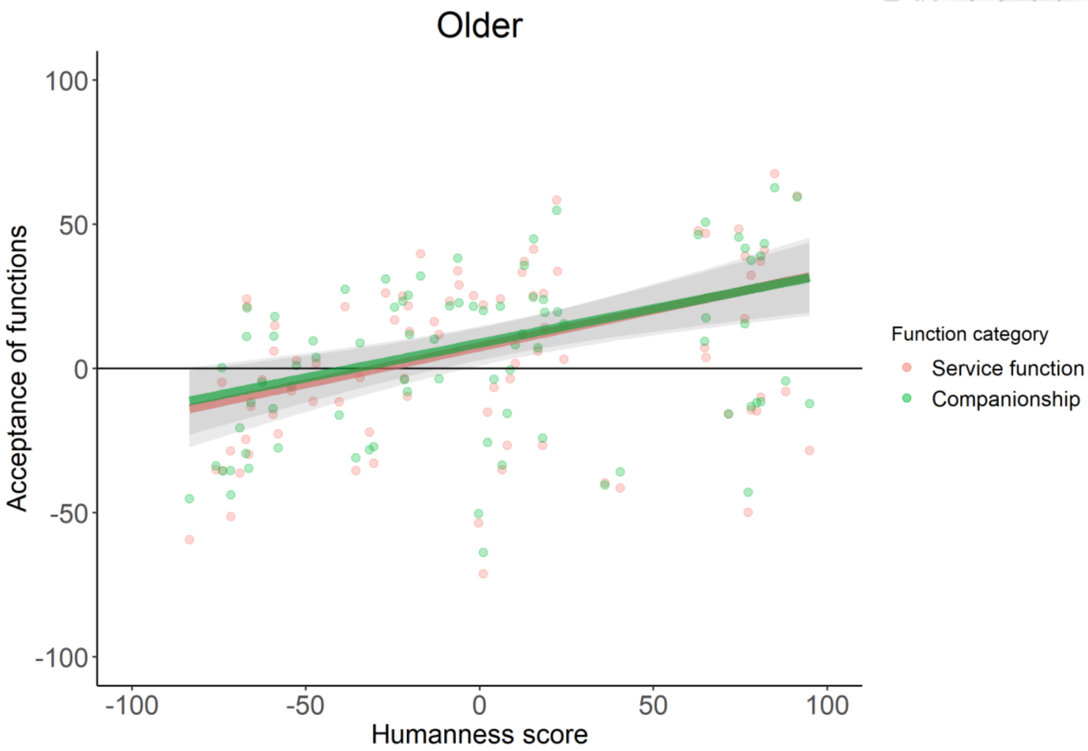

Figure 28. The comparison between acceptance of companionship and service functions from older adults

According to our data, people could also perceive different personality based on robots' appearance. The ratings from both younger and middle-aged adults showed the UVP pattern on the relation between humanness and trustworthiness/reliability, while the ratings from older adults show a linear relationship. These results indicated that both younger and middle-aged adults would trust non-humanlike and humanlike robots, but older adults only trust the humanlike robots. The ratings on friendliness/approachability from younger and middle-aged adults also showed the UVP pattern. Again, older adults perceived higher friendliness/approachability on humanlike robots. The relations between authoritativeness and humanness had the

doi:10.6342/NTU201900877

most distinct pattern. Results showed that it is very hard for users to perceive authoritativeness from robots, although results from all groups showed the linear relation between authoritativeness and humanness scores. Besides, results on adorability from both younger and middle-aged adults demonstrated the UVP pattern, but both groups of users were not easy to perceive adorability from robots. Once again, results from older adults, however, showed a linear relation, suggesting that older adults tend to perceive adorability from humanlike robots.

Robots' perceived personalities would indeed affect user acceptance of their functions. In our experiment, three personalities – trustworthiness/reliability, friendliness/approachability, as well as adorability – were positively correlated with acceptance of functions, meaning that users tended to accept assistances from robots with these positive personalities. It is possible that people would change their evaluations on robots according to perceived personalities of robots because people also evaluate other performance based on their personality. Brown et al. (2002) found that employee's personalities would affect subjective ratings on their work performance from their supervisors. Barrick and Mount (1991) also proposed that supervisors could select personnel and train or appraise employees according to their personalities. Thus, users might also judge robots' performance based on perceived personalities of robots and consequently affect the acceptance to robots. Although our

doi:10.6342/NTU201900877

questions for rating robots' personalities are not fully consistent with the big five personalities (Goldberg, 1992), we still had caught some personalities of robots. According to Goldberg (1990), extraversion factor contains gregariousness (friendliness/approachability in our study), agreeable factor contains amiability (adorability in our study), and conscientiousness factor contains dependability (trustworthiness/reliability in our study). Besides, Powers et al. (2005) used robots designed with different social cues such as appearance, voice, and demeanor to make participants be able to catch the traits of robots. They found that participants did change their interacting style with robots after detecting social cues from robots. In our data, users could also perceive robots with social cues (different personalities), and they might have different anticipations toward robots.

Apart from previous three kind of personalities, authoritativeness has multiple effects on acceptance of robots. Authoritativeness refer to whether one's opinion would be accepted because of his or her expertise (McCroskey, 1966). This concept is different from "trustworthiness/reliability", because authoritativeness contains different meaning including knowledge, profession, and status between users and robots. Thus, we predicted that people would like to be served by robots with authoritativeness because people would be willing to be served by professional workers. And they would not change their acceptance of companionship based on

doi:10.6342/NTU201900877

authoritativeness because people care whether robots could offer them social support instead of professional advices. To our surprise, only results from younger adults were in agreement with our predictions. Older adults tended to accept all kinds of assistance from authoritative robots. However, middle-aged adults preferred to be served and accompanied by robots with middle-level authoritativeness. Tay, Jung, and Park (2014) found that people would choose the robots which fit their stereotypes to help them. Perhaps users in different age groups would have stereotypes for those service and companion robots, so they have different preferences of robot appearance.

Limitations and future direction

The cohort effect (Banatvala et al., 1993), caused by sharing cultural background, social customs, values, experience, education level, and living habits in participants' growing time instead of the present study's manipulation, is one possible explanation for the observed differences of the UVP in the present study. Younger adults might have higher chances to approach robots than older adults in modern society. Younger adults grow up in the environment filled with information about new technologies (e.g., advanced robots). With these experiences, it is predictable that younger adults show lower negative feelings toward robots. However, middle-aged and older adults have less experiences to interact with this kind of new technologies. Thus, our results showed that younger adults had higher preference on non-humanlike

doi:10.6342/NTU201900877

over humanlike robots than middle-aged and older adults, which might be caused by different life experience of different aged groups. If we want to exclude this possibility, we might need to execute a cross-sequential research to ensure similar result persist in people across different generations. Nevertheless, this would not affect the fact that, regardless the underlying mechanisms, different curves were found for different age groups in terms of robot appearance.

There are some animal-like robots in our stimuli (PARO and ICAT). Although they belong to non-humanlike robots according our results, they are still different from other mechanical robots. Although robotic pets could not provide service functions, studies indicated that they could bring many benefits to users. For example, Wada and Shibata (2007) reported that the seal robot – Paro – could enhance older adults' psychological wellbeing. However, there is no UVP research on the animal-like appearance. It is also interesting to see whether cuteness is one factor dominating the UVP curve; e.g., primates are more similar to human but they might not be as likable as pets. Future studies can investigate whether the UVP exists for animal-like robots. Besides, we directly use all data to fit models without considering subjectively ratings on beauty across different appearance. Thus, features of pictures might be one possible reason for current results.

In the present study, we first tried to examine whether the UVP existed across

doi:10.6342/NTU201900877

different age groups, so we used static pictures as our stimuli. Further research with highly active robots might be needed because the movements of robots could also change the UVP (Mori, 1970). Besides, users expect robots with hands and legs to help them to process many daily lives functions (Chu et al., 2019). However, our experiment only adopted pictures containing only faces of robots. Thus, it would be also interesting to conduct studies using pictures with the full body of the robot to investigate the UVP across different age groups. In addition, voice of the robot is also critical. Because inconsistency of voice and appearance may also produce the negative feelings to robots (Mitchell et al., 2011). Robot productions should properly match robots' look and voice to maximize users' acceptance. In addition, the context could be another important feature of the UVP. To elaborate, users might have different preferences on robots' appearance in different usage scenarios such as the library and the home because they perhaps would like to use robots with trustworthiness and reliability at home while they would like to be served by robots with authoritativeness in the library.

To sum up, proper appearance for robots is important because users would change their preference and acceptance of robots based on appearance. Although we cannot clearly point out all reasons resulting these differences on acceptance of robots in the present study, we demonstrated that the design of assistive and companion

doi:10.6342/NTU201900877

robots should take UVP into consideration by customizing robots' appearances and functions according to different age groups.

doi:10.6342/NTU201900877

References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*(3), 387.

Banatvala, N., Mayo, K., Megraud, F., Jennings, R., Deeks, J., & Feldman, R. (1993). The cohort effect and Helicobacter pylori. *Journal of Infectious Diseases, 168*(1), 219-221.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1-26.

Bartneck C, Kanda T, Ishiguro H, Hagita N (2007) Is the uncanny valley an uncanny cliff? In: *16th IEEE international symposium on robot and human interactive communication*, RO-MAN 2007, Jeju, Korea, pp 368-373.

Beer, J. M., Prakash, A., Mitzner, T. L., and Rogers, W. A. (2011). *Understanding Robot Acceptance* (Technical Report no HFA-TR-1103). Atlanta, GA: Georgia Institute of Technology.

Brown, T. J., Mowen, J. C., Donavan, D. T., & Licata, J. W. (2002). The customer orientation of service workers: Personality trait effects on self-and supervisor performance ratings. *Journal of Marketing Research, 39*(1), 110-119.

Carstensen, L. L. (1995). Evidence for a life-span theory of socioemotional

doi:10.6342/NTU201900877

selectivity. *Current Directions in Psychological Science, 4*(5), 151-156.

Cheng, Y.-H., Shyi, G. C.-W., & Cheng, K.-H. (2016). Age differences in face memory and face processing between younger and older adults in Taiwan. *Chinese Journal of Psychology, 58*(4), 233-262.

Chu, L., Chen, H.-W., Cheng, P.-Y., Ho, P., Weng, I.-T., Yang, P.-L., . . . Wang, T.-M. (2019). Identifying features that enhance older adults' acceptance of robots: A mixed methods study. *Gerontology*, 1-10.

Curtis, V., Aunger, R., & Rabie, T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 271* [suppl_4], S131-S133.

Ezer, N., Fisk, A. D., & Rogers, W. A. (2009). More than a servant: Self-reported willingness of younger and older adults to having a robot perform interactive and critical tasks in the home. *Proceedings of the human factors and ergonomics society annual meeting*, 53(2), 136-140.

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., . . . Burke, G. (2001). Frailty in older adults: Evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 56*(3), M146-M157.

Goldberg, L. R. (1990). An alternative" description of personality": The big-five

doi:10.6342/NTU201900877

factor structure. *Journal of Personality and Social Psychology, 59*, 1216.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26.

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: The almere model. *International Journal of Social Robotics, 2*(4), 361-375.

Ho, C.-C., & MacDorman, K. F. (2017). Measuring the uncanny valley effect. *International Journal of Social Robotics, 9*(1), 129-139.

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of Communication, 56*, 754-772.

Lewkowicz, D. J., & Ghazanfar, A. A. (2012). The development of the uncanny valley in infants. *Developmental Psychobiology, 54*(2), 124-132.

MacDorman, K. F., & Entezari, S. O. (2015). Individual differences predict sensitivity to the uncanny valley. *Interaction Studies, 16*(2), 141-172.

MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies, 7*(3), 297-337.

Macunovich, D. J., Easterlin, R. A., Schaeffer, C. M., & Crimmins, E. M. (1995). Echoes of the baby boom and bust: Recent and prospective changes in living

doi:10.6342/NTU201900877

alone among elderly widows in the United States. *Demography, 32*(1), 17-28.

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition, 146*, 22-32.

McCroskey, J. C. (1966). Scales for the measurement of ethos. *Speech Monographs, 33*, 65-72.

Mitchell, W. J., Szerszen Sr, K. A., Lu, A. S., Schermerhorn, P. W., Scheutz, M., & MacDorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *i-Perception, 2*(1), 10-12.

Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33-35.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human performance, 11*(2-3), 145-165.

World Health Organization, National Institute on Aging, National Institutes of Health, U.S. Department of Health and Human Services. (2012). *Global Health and Aging Report.* Retrieved from: http://www.nia.nih.gov.

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human-robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics, 24*(2), 442-451.

doi:10.6342/NTU201900877

Park, S.-H., Mahony, D. F., Kim, Y., & Do Kim, Y. (2015). Curiosity generating advertisements and their impact on sport consumer behavior. *Sport Management Review, 18*, 359-369.

Powers A, Kramer ADI, Lim S, Kuo J, Lee S, Kiesler S (2005, August). *Eliciting information from people with a gendered humanoid robot.* Proceedings of the 2005 IEEE international workshop on robots and human interactive communication, Nashville, TN.

Reher, D., & Requena, M. (2018). Living alone in later life: A global perspective. *Population and Development Review, 44*, 427-454.

Republic of China, Ministry of the Interior (2019). *Number of Population by Age of 0-14, 15-64, 65 + , dependency ratio, Index of aging and dependency ratio.* Taipei, Taiwan: Ministry of the Interior.

Roy, N., Baltus, G., Fox, D., Gemperle, F., Goetz, J., Hirsch, T., …Thrun, S. (2000, January). *Towards personal service robots for the elderly*. Proceedings of the Workshop on Interactive Robotics and Entertainment (WIRE), Pittsburgh, PA.

Samochowiec, J., & Florack, A. (2010). Intercultural contact under uncertainty: The impact of predictability and anxiety on the willingness to interact with a member from an unknown cultural group. *International Journal of Intercultural Relations, 34*(5), 507-515.

doi:10.6342/NTU201900877

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2011). The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience, 7*(4), 413-422.

Seyama, J. i., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments, 16*, 337-351.

Smarr, C. A., Fausset, C. B., & Rogers, W. A. (2011). *Understanding the potential for robot assistance for older adults in the home environment* (Technical Report no HFA-TR-1102). Atlanta, GA: Georgia Institute of Technology.

Smarr, C.-A., Mitzner, T. L., Beer, J. M., Prakash, A., Chen, T. L., Kemp, C. C., & Rogers, W. A. (2014). Domestic robots for older adults: Attitudes, preferences, and potential. *International Journal of Social Robotics, 6*(2), 229-247.

Smarr, C. A., Prakash, A., Beer, J. M., Mitzner, T. L., Kemp, C. C., & Rogers, W. A. (2012, October). *Older adults' preferences for and acceptance of robot assistance for everyday living tasks.* Proceedings of the human factors and ergonomics society annual meeting, Boston, MA.

Tay, B., Jung, Y., & Park, T. (2014). When stereotypes meet robots: The double-edge sword of robot gender and personality in human–robot interaction. *Computers*

doi:10.6342/NTU201900877

in Human Behavior, 38, 75-84. doi:https://doi.org/10.1016/j.chb.2014.05.014

Tomassini, C., Glaser, K., Wolf, D. A., Broese van Groenou, M. I. & Grundy, E. (2004). Living Arrangements among Older People: An Overview of Trends in Europe and the USA*, Population Trends*, *115*, 24-34.

Tung, F. W. (2011, July). *Influence of gender and age on the attitudes of children towards humanoid robots.* International Conference on Human-Computer Interaction., Heidelberg, Germany.

Tung, F. W. (2016). Child perception of humanoid robot appearance and behavior. *International Journal of Human-Computer Interaction, 32*(6), 493-502.

Vail, K. E., Rothschild, Z. K., Weise, D. R., Solomon, S., Pyszczynski, T., & Greenberg, J. (2010). A terror management analysis of the psychological functions of religion. *Personality and Social Psychology Review, 14*(1), 84-94.

Vess, M., Arndt, J., & Cox, C. R. (2012). Faith and nature: The effect of death-relevant cognitions on the relationship between religious fundamentalism and connectedness to nature. *Social Psychological and Personality Science, 3*(3), 333-340.

Wada, K., & Shibata, T. (2007). Living with seal robots—its sociopsychological and physiological influences on the elderly at a care house. *IEEE Transactions on Robotics, 23*, 972-980.

doi:10.6342/NTU201900877

Yeh, S.-C., & Lo, S. K. (2004). Living alone, social support, and feeling lonely among the elderly. *Social Behavior and Personality: An International Journal, 32*(2), 129-138.

doi:10.6342/NTU201900877

Appendices

Appendix. A

問卷	題項
好奇心問卷	1. “您對於機器人有多好奇？”
	2. “您願意花時間看機器人的介紹影片嗎？”
	3. “您願意多了解機器人嗎？”
	4. “如果容易取得機器人的使用介紹，您願意去看這些資料嗎？”
	5. “您願意主動上網去找關於機器人的介紹嗎？”
	6. “如果您家附近的活動中心有機器人的進階教學課程，您願意主動參加嗎？”

doi:10.6342/NTU201900877

Appendix. B

問題類型	指標	題項
服務功能	日常功能	1. 當我無法自理時，我願意讓這台機器人協助我的生活起居(包括家事、清潔、準備餐點、和遞東西等功能)。
	保護	2. 我願意讓這台機器人保護我。
	觸碰	3. 我願意被這台機器人碰觸(包括協助走動、按摩、和攙扶起身等功能)。
	隱私	4. 我願意讓這台機器人協助我處理涉及個人隱私的事（包括洗澡、穿衣、餵食、和管理財務等功能)。
陪伴功能	陪伴	5. 我願意讓這台機器人陪在我身邊。
	娛樂	6. 我願意跟這台機器人一起休閒娛樂。
	聊天	7. 我跟這台機器人可以像朋友般地聊天。
	分享心事	8. 我願意和這台機器人分享心事。
恐怖谷	喜歡	9. 我 _____ 這台機器人 不喜歡 – 喜歡
	厭惡	10. 我覺得這台機器人 _____. 非常不噁心 – 非常噁心
性格	信賴且可靠	11. 這台機器人看起來 _____ 非常不可靠、不值得信賴 – 非常可靠、值得信賴
	友善及可親近	12. 這台機器人看起來 _____ 非常不友善且難親近 – 非常友善且容易親近
	權威程度	13. 這台機器人看起來 _____ 非常沒有權威 –非常有權威
	可愛程度	14. 這台機器人看起來 _____ 非常不可愛 –非常可愛
	性別	15. 這台機器人的性別是？ 男– 無法區辨 – 女
恐怖谷	像人程度	16. 這台機器人看起來. 非常不像人 – 非常像人
檢驗題		17. 我以前看過這台機器人。 是 / 否 / 無法判斷

doi:10.6342/NTU201900877