Attention Distribution and Decision-Making in the Process of Robot's Appearance Design and Selection

Nicholas Hong Li Khoo[1], Fan Li[2(✉)], Chun-Hsien Chen[1], Yisi Liu[3], Fitri Trapsilawati[4], and Olga Sourina[5]

[1] School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore, Singapore
[2] Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic, University, Hung Hom, Hongkong
fan-5.li@polyu.edu.hk
[3] Centre of Excellence in Maritime Safety (CEMS), Singapore Polytechnic, Singapore, Singapore
[4] Universitas Gadjah Mada, Yogyakarta, Indonesia
[5] Fraunhofer Singapore, Nanyang Technological University, Singapore, Singapore

Abstract. Humanoid robots gaining popularity in the service industry all over the world. When designing humanoid robots, product designers put additional thought into the aesthetic design to avoid the complications that arise from the "uncanny valley". It has long been found that genders, personalities, and cultural upbringings may affect individual ratings on the appearance of humanoid robots. To better aid and hasten the appearance design and selection process, this report explores the use of Electroencephalography (EEG) and vision tracking devices to evaluate preferences based on "aesthetic" criteria of robot designs. Based on the analysis, it was found that individuals, regardless of the three traits, paid more attention to their top choices during the selection. In rating the robots' appearance, it was also found that Eastern students rated a significantly higher score of "intelligence" than Western students. However, both eastern and western students have a similar choice of their favorite robots.

Keywords: Electroencephalography · Eye-tracking · Robots appearance design · Preference

1 Introduction

Recently, the service industry has seen an increase in the adoption of robots, such as robots for elderly people, robots for children, and robots for patients [1, 2]. For example, Pepper Parlour Café adopted robots for taking customer orders and cleaning up. With the increase in the adoption of robots in service industries, their design has been increasingly important, as service industries are normally customer-centered markets. Unlike industrial robots, which are created for the sole purpose of performing repetitive

P.-L. P. Rau (Ed.): HCII 2022, LNCS 13314, pp. 535–544, 2022.
https://doi.org/10.1007/978-3-031-06053-3_36

processes, requiring less focus on their aesthetic design, the service robots design is much more complicated due to its requirements for human enjoyment [3].

According to the uncanny valley of the humanoid robot [4], a robot with a bad social design may be off-putting instead of engaging. Hence, a number of studies have been conducted to test the human perception of the appearance design of service robots. It has been found that both physical aesthetic features and emotional preference features are important influencing factors of human perception of service robots [5]. In addition, to maximize user experience in human-robot interactions, designers should try to match the appearance with communication [2]. Beyond the physical design of a robot's appearance, an individual's personality, cultural difference, and gender contribute to the difference in perception of robots [6]. A study recruited subjects from China, Korea, and Germany to investigate their response to robots appearance. The study found that subjects from low-context cultures showed less engagement to the less sociable robots [7]. The gender effects on the perception of users of robots appearance are significant, too. Males seem to have a relatively more positive attitude toward robots in healthcare than females. Normally, the response to robots' appearance, such as task engagements and attitudes can be assessed [7]. The subjective questionnaire, such as quality of service can be used, too. Nevertheless, limited studies investigated the visual attention distribution and decision processing during the evaluation of robots appearance. In this study, a novel method based on electroencephalography (EEG) and eye-tracking devices has been adopted for understanding the evaluation process. It is expected the novel approach can assist companies to design and develop an objective approach to assess the response of consumers to service robots [4]. These psychophysiological data generated from consumers can be transformed into meaningful design recommendations and deliver more tailored, customer-centric services robots.

The objectives of this project are to evaluate the effect of different cultural backgrounds (Eastern vs. Western) on an individual's selection of humanoid service robot design. The organization of this study is as follows: Sect. 2 reviews the effects of cultural differences on users' perception of service robots. Section 3 lays out the methodologies and tools utilized in the experiment and the recorded variables. Section 4 elaborates on the key findings, experimental limitations and the proposed design for a humanoid robot to be adopted in the service industry. Section 5 delivers the conclusion of the study and recommendations for future research.

2 Literature Review

2.1 Cultural Differences Bias

Different cultures have led to different traditions for robots. The positive attitude towards robots in Eastern culture, especially the Japanese, is largely based on the image conveyed by popular media such as movies and Anime cartoons [8]. On the other hand, attitudes toward the Western culture are less positive because of the destructive and hostile robots portrayed in movies such as "Terminator Genisys" or "The Matrix".

The Western world believes that there should be a separation between nature and culture, to organize the world in an orderly and precise fashion. There is no place for hybridization in such classifications [9]. In Japan, all intermediary beings are part of the big picture. As shown in Fig. 1, the Japanese make no distinction between them and try to create links to form a continuous network of beings [6].

The Western world views robots as a step backward on a social level and that it creates paranoia of replacing people. The Japanese choose to adopt a different belief by viewing robots as a means of turning away from the digital virtual world to more human-like objects, thus facilitating contact in the physical world again [6].

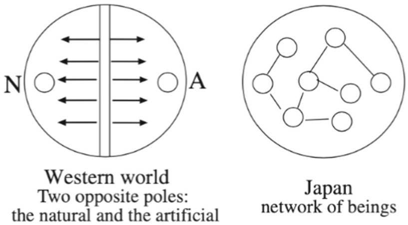

Fig. 1. Distinguishing vs Linking Beings [6]

2.2 Gender Bias

Existing research has illustrated that the simple gendering of robots through manipulation of their voice and name can affect an individuals' behavior towards it. However, this is also dependent on other factors such as the evaluator's gender [10].

The preference of robots by either gender is a key study in HRI and determines the amount of service received by the robot. Nomura et al. mentioned in their research that females have more negative feelings towards the interaction with robots compared to males [10]. This can be supported by an experiment conducted by Kuo et al., where humans interacted with a robot in a healthcare environment. It was found that females were less receptive to the usefulness of the robot and the possibility of using it in the future [11].

Eyssel and Hegel investigated the effects of facial cue manipulation and how it relates to human gender stereotypes for traits and tasks [12]. By altering simple facial features such as adding short hair (male robot) or long hair (female robot), they revealed that male robots were perceived as being more masculine than female robots, and stereotypically male tasks were perceived as more suitable for male robots.

2.3 Myers-Briggs Type Indicator (MBTI)

A personality is defined as the combination of qualities or characteristics that form an individual's distinctive character, usually evolved from biological or environmental factors. Individuals with different personalities behave differently and this influences their economic choices [13] such as willingness to pay for a product as well as career choices [14]. Personalities are also capable of influencing an individual's perception of an object, which generally involves complex psychological mechanisms and processes in the perceiver.

To help psychologists better understand the different personality types, one popular "sorter" is the Myers-Briggs Type Indicator (MBTI), devised by Isabel Briggs Myer and her mother, Katharine Briggs. The theory of MBTI comes from Swiss psychiatrist, Carl G. Jung, who observed that people engaged in one of two mental functions: perceiving or judging. Today, the two main applications of MBTI are to identify basic preferences of each of the four dichotomies specified in Jung's theory and identify the 16 distinctive personality types that result from the interactions among the preferences [14].

Based on Myers Briggs theory, there are four preference types [15]: Sensing (S) and Intuition (N) determines if the individual pays more attention to the information derived from their five senses (Sensing) or to the patterns in the information in which they receive (Intuition). Individuals with "S" traits tend to be practical while those with "N" traits exhibit traits similar to that of a Visionary [15].

3 Methodology

To test the effects of robots appearance on the perception of consumers, five steps, including robots' appearance design, area of interest (AOI) generation, user experiments, data collection, and data analytics have been conducted, as shown in Fig. 2.

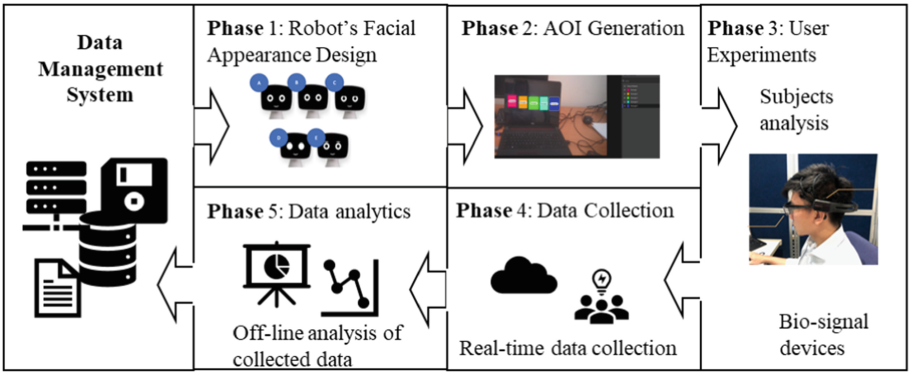

Fig. 2. A general framework work of the study

3.1 Robots Facial Appearance Design

Several preliminary studies have been conducted, it was identified that human perception of service robots is largely influenced by the upper body appearance. This comes as no surprise as humans commonly form impressions of others' faces and also tend to do so with artificial agents. With past research having a strong focus on robots' upper body appearance, especially their facial appearance, the experiment conducted focuses on the robots' facial features. In this study, participants are shown 5 controlled robot images, namely baseline, no mouth, eyelids, no eyes, and eyebrows, as shown in Fig. 3. The images differ by one feature from the base image (Image A) to increase the difficulty of robot design selection.

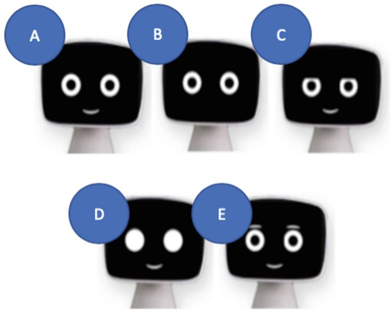

Fig. 3. Robot images

3.2 AOI Generation

This study focuses on figuring out the visual distribution and the perception of users in evaluating the appearance of robots. Hence, all the five robots are presented to subjects on one page. To investigate the visual distribution, which is represented by fixation number and fixation duration, each face of the robot is selected as one AOI. Five rectangles of the same size are designed. The rectangle can cover the robot face with a 1mm allowance on both length and width.

3.3 User Experiments

The study has been approved by the Institutional Review Board of NTU (IRB-2019-02-032). 30 participants (17 Females and 13 Males) between the age of 21–24 years were recruited for the experiment. Of the 17 females, 11 originated from the Eastern world (nations in Asia including the Middle East) while 6 originated from the Western world (North and South America, Europe, Australia, and New Zealand). Of the 13 males, 8 originated from the Eastern world and 5 originated from the Western world.

The device to be used in the experiment is the EMOTIV EPOC+. The device is equipped with 14 channels with a sample rate of 2048. The eye tracker used in the experiment is the Tobii Pro Glasses 2, which has a one-point calibration procedure for easy setup.

The entire experiment includes 3 stages, as shown in Fig. 4. In the first stage, participants are briefed on the flow as well as objectives of the experiment to give them a better understanding of how the recorded data is relevant to the study. In the second stage, both EEG and Eye Tracker are calibrated for every new participant to minimize discrepancy in data. To ensure accuracy in data recorded, a 100% fit rating has to be detected by the device before proceeding with the experiment. To calibrate the eye tracker, participants are positioned at a comfortable distance away from the computer screen (50cm). The eye-tracker captures the calibration spots on the screen before commencing the actual experiment. In the last stage, the images of robot faces are shown to participants, and questionnaires are designed to evaluate their responses.

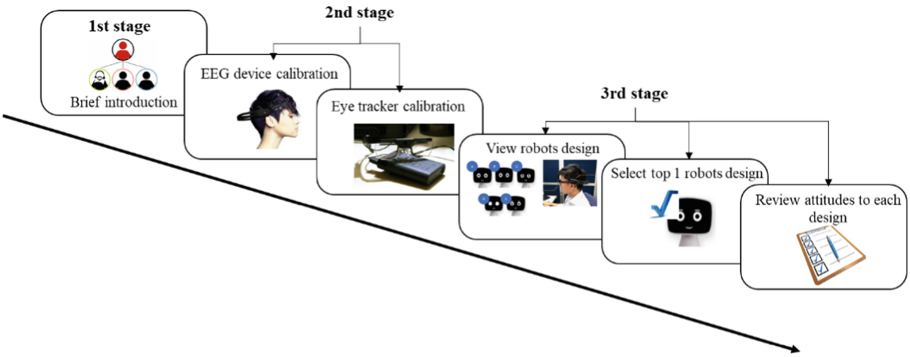

Fig. 4. The process of user experiments

3.4 Data Collection

Two types of data were collected, namely subjective report data via questionnaire and objective biosignals via Emotive and Tobii glass. The subjective responses to the robots appearance were collected after user experiments with personal metric ratings, likability, friendliness, intelligence, valence, arousal, and dominance. The personal metric ratings are derived from standardized models. This experiment uses Arousal, Dominance, and Valence as evaluation metrics, proposed by Russel and Mehrabian [16]. Likability, smartness, and friendliness are subsets for Human-Robot Interaction (HRI) [17]. Raw EEG data and eye-tracking data were collected during the user experiments. The EEG device is capable of recording the EEG signals at 120 Hz. The Tobii glass has a sampling rate of 100 Hz.

3.5 Data Analytics

Figure 5. shows the process of data analysis. Emotions values, namely positive and negative were obtained via EEG data analytics. To evaluate the raw EEG data, an EEG Subject-independent Processor developed by Fraunhofer Singapore was used. The EEG Subject-independent Processor is a stand-alone off-line application that uses advanced transfer learning algorithms to recognize different physiological signals from raw EEG data [18, 19]. The software is able to recognize the mental states of a subject without the need for any calibration. This is achieved through the use of an existing database consisting of EEG data with labels from other subjects. The window is sliding over the raw EEG data with 1-s step. The first sample is extracted from 0–4 s on the raw data, and the second example is extracted from 1–5 s on the raw data. The process repeats until the window reaches the end of the whole signal. Positive emotion is indicated with a “1” and negative emotion is represented with a “0”. The extracted emotions are then labeled with “P” for positive and “N” for negative.

The Time To First Fixation (TTFF), fixation durations, and fixation counts were obtained from eye-tracking data. The I-VT method [20] was applied in the raw eye-tracking data to extract the fixations based on gaze velocity. According to the fixation position and AOI position, the TTFF, fixation durations, and fixation counts were calculated for each AOI. Analysis of variance (ANOVA) was carried out to determine if there is a correlation between the individual's top choice and their vision tracking data. Biserial Correlation was used to analyze all the collected data.

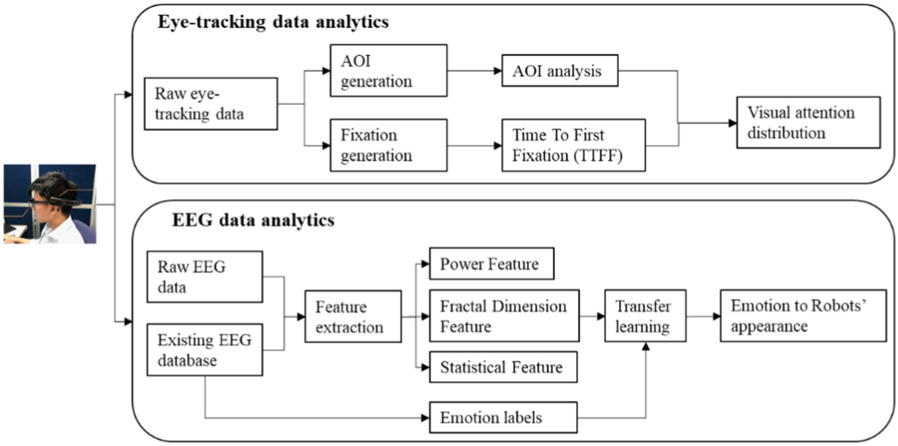

Fig. 5. EEG and Eye movement-based method for appearance evaluation

4 Results and Discussion

The responses gathered from participants were first analyzed using Biserial correlation analysis. The analysis is used to determine if there is any correlation between the likability, friendliness, intelligence, valence, arousal, and dominance of the robot and gender, culture, and personality of the respondent. A *p*-value below 0.05 indicates that the two variables are correlated.

Based on the data analysis, the following conclusions can be derived:

- Individuals' gender is correlated with the perceived intelligence, valence, and dominance of a robot.
- Individuals' culture is correlated with the perceived friendliness and dominance of a robot.
- Individuals' personality is correlated with the perceived intelligence and valence of a robot.

ANOVA was conducted to test the effects of gender, culture, and personality on users' perceptions of the humanoid robot design. The ANOVA results showed that gender and culture have significant effects on users' perception of the intelligence of the robot design. Specifically, males felt the Tapia is intelligent while females felt it was normal. Though the effects of personality on user's perception are not significant, the interaction effects between personality and gender were significant (intelligence $F_{(1,26)} = 7.35$, $p < 0.05$; valence: $F_{(1,26)} = 5.92$, $p < 0.05$; arousal: $F_{(1,26)} = 10.233$, $p < 0.05$).

The participants' most liked and disliked robot designs were identified based on the questionnaire and matched with their "TTFF", "Fixation Durations" and "Visit Counts".

Analysis of variance (ANOVA) was carried out to determine if there is a correlation between the individual's top choice and their vision tracking data. For Time To First Fixation (TTFF), $F_{(1,26)} = 0.56$, $p = 0.07$. For Fixation Duration, $F_{(1,26)} = 10.18$, $p < 0.05$. For visit counts, $F_{(1,26)} = 1.38$, $p = 0.2$. Based on the results gathered, only "Fixation Duration" has a p-value less than 0.05. As a result, it can be concluded that participants tend to spend more time fixated on their top choice of robot design.

Biserial Correlation was used to analyze the EEG-based emotion data set and it was shown that emotions recorded were not consistent with their choices. However, this could be attributed to the noise in emotion data.

5 Results and Discussion

After an arduous product design process, product designers have to decide on which product is best to launch to market. As the service industry sees a rise in the adoption of robots to replace manual labor, robotic companies struggle to design a suitable robot that avoids the effects of the "Uncanny Valley". When it comes to the design of robots. which require social interaction with humans, the design process becomes increasingly complex as it requires the consideration of representation, anthropomorphism, and task orientation, depending on the service context. The report covered several important attributes to be examined in robot design such as its perceived friendliness and intelligence. It also explored how one's gender, personality, and cultural difference influences their perception of a product. In the case of the robot design study, results showed that gender and culture are closely correlated to users' perception of intelligence. The interaction effects between personality and gender were also significant. In addition, it was also observed that the amount of time spent looking at a particular design further confirms the participants' choice of the preferred design. For EEG metrics, although no significant findings were derived, proper calibration using IADS will bring about new findings in the field of product design selection. The combination of both EEG and vision-tracking devices, would ultimately accelerate the product design selection process and ease the mental workload of product designers.

6 Appendix D – Computer Experiment Questionnaire

6.1 Test Scenario

Group Images.
Pick your 1st, 2nd, and 3rd choice of design out of the whole set. Please give a brief explanation as to why those three were your preferred design.

Individual Images.
Utilizing the 9-Point Hedonic Scale:

1. Likability is closely related to a positive emotion. Rate the likability of the robot image that you have just seen.
2. Friendliness is an important factor to be considered for robots employed in the service industry due to its interaction with humans. Rate the friendliness of the robot image that you have just seen.
3. Smartness determines how useful the robot is perceived to be and gives consumers the confidence to approach the robots for assistance. Rate the smartness of the robot image that you have just seen.
4. Valence is defined as the feeling of pleasure derived from the robot. Rate the valence level.
5. Arousal is the level of attention drawn by the robot. Rate the arousal level.
7. Dominance is defined as an overwhelming feeling induced by the robot. Rate the dominance level.

Acknowledgment. Thanks for all the subjects that participated in the user experiments.

References

1. Walters, M.L., Koay, K.L., Syrdal, D.S., Dautenhahn, K., Te Boekhorst, R.: Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials. Procs of New Frontiers in Human-Robot Interaction (2009)
2. Klüber, K., Onnasch, L.: Appearance is not everything-preferred feature combinations for care robots. Comput. Hum. Behav. **128**, 107128 (2022)
3. Pokojski, J.: Evaluation of humanoid robot design based on global eye-tracking metrics. in transdisciplinary engineering for complex socio-technical systems–real-life applications. In: Proceedings of the 27th ISTE International Conference on Transdisciplinary Engineering, July 1–July 10, 2020. IOS Press (2020)
4. Mori, M., MacDorman, K.F., Kageki, N.: The uncanny valley [from the field]. IEEE Robot. Autom. Mag. **19**(2), 98–100 (2012)
5. Gan, Y., et al.: Integrating aesthetic and emotional preferences in social robot design: an affective design approach with Kansei engineering and deep convolutional generative adversarial network. Int. J. Ind. Ergon. **83**, 103128 (2021)
6. Kaplan, F.: Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. Int. J. Humanoid Rob. **1**(03), 465–480 (2004)
7. Li, D., Rau, P.-L., Li, Y.: A cross-cultural study: effect of robot appearance and task. Int. J. Soc. Robot. **2**(2), 175–186 (2010)
8. Bartneck, C., Hu, J.: Rapid prototyping for interactive robots (2004)
9. Latour, B.: We Have Never Been Modern. Harvard University Press, Cambridge (2012)
10. Nomura, T.: Robots and gender. Gender and the Genome **1**(1), 18–25 (2017)
11. Kuo, I.H., et al.: Age and gender factors in user acceptance of healthcare robots. In: RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication. IEEE (2009)
12. Eyssel, F., Hegel, F.: (s) he's got the look: Gender stereotyping of robots 1. J. Appl. Soc. Psychol. **42**(9), 2213–2230 (2012)
13. Boyce, C., Czajkowski, M., Hanley, N.: Personality and economic choices. J. Environ. Econ. Manag. **94**, 82–100 (2019)

14. Varvel, T., S.G. Adams, and S.J. Pridie. A study of the effect of the Myers-Briggs Type Indicator on team effectiveness. In: Proceedings of the 2003 ASEE Annual Conference and Exposition, Session (2003)
15. Quenk, N.L.: Essentials of Myers-Briggs type indicator assessment. Wiley, New York (2009)
16. Russell, J.A.: Affective space is bipolar. J. Pers. Soc. Psychol. **37**(3), 345 (1979)
17. Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int. J. Soc. Robot. **1**(1), 71–81 (2009)
18. Liu, Y., et al.: Human factors evaluation of ATC operational procedures in relation to use of 3D display. In: Stanton, N. (ed.) AHFE 2019. AISC, vol. 964, pp. 715–726. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-20503-4_64
19. Liu, Y., et al.: Detection of humanoid robot design preferences using EEG and eye tracker. In: 2019 International Conference on Cyberworlds (CW). IEEE (2019)
20. Li, F., Chen, C.-H., Xu, G., Khoo, L.-P.: Hierarchical eye-tracking data analytics for human fatigue detection at a traffic control center. IEEE Trans. Hum.-Mach. Syst. **50**(5), 465–474 (2020)