

Influences of Gender Values into Interaction with Agents: An Experiment Using a Small-Sized Robot

Tatsuya Nomura^{1,2} and Kentaro Hayata¹

¹Media Informatics, Ryukoku University, Japan

²ATR Intelligent Robotics and Communication Laboratories, Japan

Abstract: To investigate the influences of humans' gender values in Human-Agent Interaction, the research conducted an experiment using a small-sized robot having no explicit gender cue (Yorisoi Ifbot) and a psychological scale measuring some types of gender values (the M-H-F scale). As a result, it was found: (1) participants' gender attribution to the agent was related only with their active impression of the agent negatively, which had no correlation with their gender values, (2) although participants' passive impression of the agent did not influence their gender attribution to the agent, it was related with their values of masculinity and femininity, which were strongly correlated each other, and (3) participants' value of masculinity and gender attribution affected their behaviors such as their time spent for interaction and numbers of smiles during interaction with the agent.

1 Introduction

Gender is one of the most important factors that have been studied for a long time in psychology. The recent research of human-agent interaction (HAI) has also focused on this factor. Some studies suggested the effects of gender assignment to agents (gendering agents) on humans [1]-[3]. Other studies found differences due to human gender on the feelings and behaviors toward agents [4], and interaction effects between human gender and other factors such as gender assigned to agents [5]-[7]. These studies suggested the possibility that gendering agents makes human-agent interaction more efficient.

On the other hand, some researchers of social sciences critically argued that gendering agents was done by technologists' arbitrary choice based on their common senses [8] and had the possibility of reproduction of gender stereotypes [9][10]. Some researchers on agents suggested that it may cause users' abusive behaviors toward agents [2]. However, these studies did sufficiently not take into account processes of humans' gender attribution to artificial agents.

Most of the existing studies on gender in HAI explicitly assigned physical gender cues to agents, such as appearance and voices, and then investigated how gender of the agents perceived by subjects affected the

subjects' feelings and behaviors. Thus, it was sufficiently not considered what factor made humans attribute gender to artificial agents. On the other hand, Nomura and Takagi [11] performed gender assignment to an agent having no physical gender cue (a small-sized humanoid robot with mechanical appearance) only by naming with gender values, and suggested that even this minimum gender assignment could evoke humans' gender impressions of the agent, and the attributed gender could have interaction effects with other factors. This study suggests the possibility that humans' gender values may cause gender perception for agents and influence humans' feelings and behaviors even if no explicit gender cue is assigned to the agents.

To investigate the influences of humans' gender values in HAI, the research conducted an experiment using a small-sized robot having no explicit gender cue and a psychological scale measuring some types of gender values. The paper reports the results and discusses about the implications.

2 Method

In the experiment, the interaction between humans and an agent was conducted in a laboratory setting to explore the influences of the humans' gender values into impressions of and behaviors toward the agent. In the



Figure 1. Yorisoi Ifbot

interaction setting, no specific context was introduced. All the procedures were conducted with the Japanese.

2.1 Date and Participants

The experiment was conducted from November to December, 2012. A total of twenty-six Japanese university students in the western part of Japan participated to the study (male: 8, female: 18, natural science and technology education: 4, social science education: 22, mean age: 20.2). The participants were not notified in advance about any concrete content of the study. They received one-thousand yen as compensation.

2.2 Agent Used in the Study

The experiment used a small-sized autonomous robot

“Yorisoi Ifbot” [12] shown in Figure 1. It has a height of 44.5 cm, a weight of 8.1 kg, a depth of 35.8 cm, and a width of 43 cm. Although the robot has no function to move around or move the arms, it has voice recognition and utterance systems, and shows the emotional expression by moving the eyes with the eyelids and the head, and lighting the LEDs on the face.

The robot can autonomously react for users’ utterances based on built-in programs. It was originally developed for the purpose of relieving the solitary feelings or loneliness of the elderly through simple conversation. The experiment did not adopt Wizard-of-Oz method but used the built-in programs. The robot had no gender cue on the appearance, voice, or contents of utterances.

To encourage participants to interact with the robot in the unstructured way, a document was prepared in advance, in which the following items related to the robot’s functions were written:

- Participants need to utter toward the robot when the LEDs on the head are switched on and off.
- Participants need to utter toward the robot clearly and a little largely.

The above items correspond to the restriction of the voice recognition function. Moreover, the document included many samples of utterances from users to the robot (e.g., “hello”, “it is raining”, “thank you”, “you are pretty”).

2.3 Procedure

Each session in the experiment was conducted in the following way:

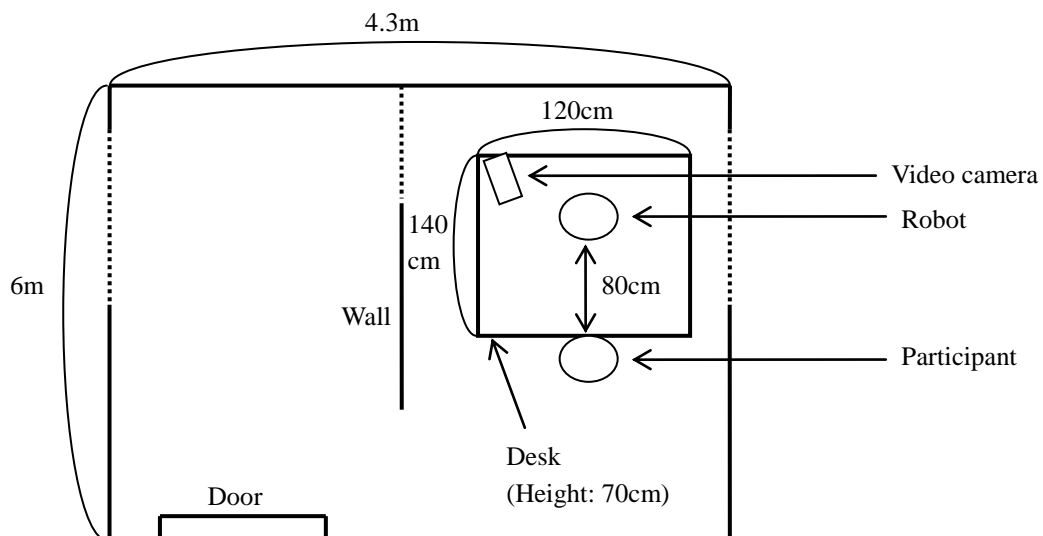


Figure 2. Overview of the room where the experiment was executed (a view from above)

1. Each participant was briefly told about the experiment and signed the consent form including the permission of video recording. In this stage, the experimenters only indicated that the task in the experiment was interaction with a robot. Then, she/he filled in a pre-session questionnaire.
2. The participant was led to an experimental room in which the robot was placed on a desk, as shown in Figure 2. The experimenters instructed the participant to sit on the chair in front of the desk, and interact with the robot freely for about five minutes. She/He was allowed to stop the interaction when she/he wished, even if five minutes did not pass.
3. The participant was left alone in the room, and started the interaction with the robot, referring to the document mentioned above.
4. The session was finished when the participant stopped the interaction and exited from the room. In case that the participant kept interacting with the robot for more than five minutes, the experimenters entered the room again and indicated that the session was finished, while deciding a break of the interaction.
5. Finally, the participant was led to the first room again, and responded to a post-session questionnaire. Then, the experimenters interviewed the subject about the robot and the experiment, and conducted debriefing about the actual aim of the experiment.

2.4 Measures

The measures in the experiment were based on self-reports from the questionnaires and video-recorded data.

The pre-session questionnaire consisted of demographic items (age, gender, educational background) and a psychological scale measuring participants' gender values, the M-H-F scale [13]. This Japanese scale consists of several adjective items representing masculinity, femininity, or humanity independent on gender. Respondents are asked to answer what degree they consider the characteristic represented by each item important for them, based on seven-graded interval (0: Not important at all – 3: Not decided – 6: Absolutely important). The experiment adopted the two subscales measuring the value of masculinity (10 items:

e.g., “muscular”, “energetic”, “bold”) and that of femininity (10 items: e.g., “pretty”, “elegant”, “slender”).

The post-session questionnaire was aimed at measuring participants' attribution of gender to and impression of the robot. Firstly, respondents were asked to answer which of male and female they attributed to the robot. Then, they were provided with twenty adjectives to measure their impression of the robots (e.g., “feminine”, “masculine”, “active”, “mild”, “gentle”). These adjectives were selected from the previous experiment [11]. Participants were asked to respond to the degrees to which they felt the impression of the robot was represented by each adjective. Each item had a score for rating with seven intervals (1: I strongly disagree – 4: Not decidable – 7: I strongly agree).

Moreover, the following behavioral indices were extracted from the video data: (1) The time that each participant spent for interaction with the robot (sec), and (2) the number that each participant smiled during the interaction. Index (1) was defined as the time spent from the experimenters' exiting from the room to the decision of finishing the session by the participant or experimenters. Index (2) was firstly identified by two persons independently, and finally decided by discussion between the two persons in cases of different identification between them.

3 Results

3.1 Gender Values and Impressions

One item of the masculinity value subscale of the M-H-F scale was lacked in many participants due to mistakes of the questionnaire layout. Since Cronbach's reliability coefficient α for the remaining 9 items was .714 showing sufficient internal consistency, these 9 items were used for measuring participants' values of masculinity. Cronbach's reliability coefficient α for the femininity value subscale was .895 and showed sufficient internal consistency. Each subscale score was calculated as the sum of the scores of the corresponding items (value of masculinity: max 54, min 0; value of femininity: max 60, min 0). The mean scores were 37.8 (*SD*: 5.0) and 35.2 (*SD*: 8.0) for values of masculinity and femininity, respectively.

To measure participants' impressions of the robot, selection of suitable adjective items was performed as follows. Although factor analysis should have been conducted for 20 adjective items, the number of the

Table 1. Adjective items for measuring impressions of the robot and internal consistencies

| | Adjective Items | Cronbach's α |
|--------------------|--|---------------------|
| Passive Impression | Bashful, Warm-hearted, Feminine, Mild, Favorable, Polite in the way of speaking, Gentle | .810 |
| Active Impression | Masculine, Active, Not content, Ambitious | .637 |

Table 2. Pearson's correlation coefficients r between gender values and impression scores

| | Feminity value | Passive impression | Active impression |
|--------------------|----------------|--------------------|-------------------|
| Masculinity value | .657*** | .514** | .110 |
| Feminity value | | .336 [†] | .300 |
| Passive impression | | | .089 |

([†] $p < .1$, ** $p < .01$, *** $p < .001$)

participants in the experiment was insufficient for the analysis. Thus, these adjective items were manually classified into two categories, and then item analyses using Cronbach's α and I-T correlations were conducted for the two item groups. The first group consisted of 11 adjectives representing passive impressions including "obedient" and "bashful". The second group consisted of 9 adjectives representing active impressions including "active" and "ambitious". The items "feminine" and "masculine" were included in the first and second groups, respectively.

On the first group, 4 items were reduced as their inclusion showed lower internal consistency. On the second group, 5 items were reduced. Table 1 shows the final impression adjective groups and Cronbach's reliability coefficients for them. Each impression score was calculated as the sum of the scores of the corresponding items (passive impression: max 35, min 7; active impression: max 20, min 4). The mean scores were 26.2 (SD : 4.2) and 11.0 (SD : 2.6) for passive and active impression, respectively.

Table 2 shows Pearson's correlation coefficients r between the gender value and impression scores. There was a strong correlation between the values of masculinity and feminity. The gender values had no

statistically significant correlations with the active impression, although they had a moderate level of correlation with the passive impression. Moreover, there was no correlation between the passive and active impressions.

3.2 Gender Attribution to the Agent

Among 8 male and 18 female participants, 6 male and 12 female participants attributed female gender to the robot. Chi-square test found no relationship between participants' gender and gender attribution to the robot ($\chi^2(1) = .181, n.s.$). The estimated interval of the rate of female gender attribution to the robot was [.482, .857] with 95% confidence.

To investigate relationships of the gender values and impressions of the robot with the gender attribution to the robot (male = 0, female = 1), a logistic regression analysis was conducted by using the backward elimination method. The extracted model (-2Log Likelihood: 19.277, Cox & Snell R^2 : .389, Nagelkerke R^2 : .549) found only the negative path from the active impression to the attribution of female gender to the robot ($B = -1.351, Wald = 4.027, p < .05$).

Table 3. Linear regression models for behavior indices

| | Number of smiles during interaction $F(1,22) = 10.761^{**}$ | | Time spent for interaction $F(2, 21) = 7.869^{**}$ | |
|--------------------|--|----------|---|---------|
| | β | t | β | t |
| Masculinity value | -.573 | -3.280** | -.418 | -2.481* |
| Gender attribution | | | -.426 | -2.530* |
| R^2 : | .298 | | .374 | |

(* $p < .05$, ** $p < .01$)

3.3 Relationships with Behaviors

The behavioral data in two participants was lacked due to a failure of the video. Thus, the behavior indices of twenty-four participants were analyzed. On the time that participants spent for interaction with the robot, the mean and standard deviation (sec) were 269.2 and 62.4, respectively. On the numbers that participants smiled during the interaction, the mean and standard deviation were 7.0 and 6.5, respectively.

To investigate relationships of the gender values, impressions of, and gender attribution to the robot (male = 0, female = 1) with behavioral indices, linear regression analyses were conducted by using the backward elimination method. Table 3 shows the extracted models. The masculinity values negatively affected both the time spent for and the numbers of smiles during interaction with the robot. The female gender attribution with the robot also negatively affected the time spent for interaction with the robot.

4 Discussion

4.1 Implications

The agent used in the experiment did not show extreme bias of the participants' gender attribution. The participants' gender attribution was related only with their active impression of the agent, which had no correlation with their gender values. Although the passive impression did not influence the gender attribution to the agent, it was related with the values of masculinity and femininity, which were strongly correlated each other. On the other hand, the participants' value of masculinity and gender attribution affected their behaviors such as the time for interaction and the numbers of smiles.

These results suggest the following facts. Even if agents do not have concrete gender cues, users may have some impressions of the agents, while referring to their gender values. Although users' gender values do directly not affect their gender attribution to the agents, their impressions may influence the gender attribution, and their gender values and attributed gender may affect the behaviors. To sum up, although de-gendering in agent technologies may reduce the reproduction of gender stereotypes, it may not solve behavioral differences based on gender values users originally have.

4.2 Limitations

The experiment was based on a specific type of agent, a small number of participants, and a single culture of participants. Moreover, the measurement of gender values and impressions of the agent may not have had sufficient validity due to the lack of data and the small number of samples. Thus, the results cannot be generalized for other types of agents and users having other cultures, while it can provide with a possibility. In particular, gender values themselves depend on cultures and times. In this sense, the M-H-F scale used in the experiment may not be suitable for the purpose.

Future works should be extended to a variety of users and agents, using more sophisticated measure of gender values.

Acknowledgement

The research was supported in part by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research No. 21118006.

References

- [1] A.I. Niculescu, D.H.W. Hofs, E.M.A.G. Van Dijk, and A. Nijholt, "How the agent's gender influence users' evaluation of a QA system," in *Int. Conf. User Science and Engineering (iUSER)*, 2010, pp. 16-20.
- [2] A. De Angeli and S. Brahnam, "Sex stereotypes and conversational agents," in *Ws. Gender and Interaction: Real and Virtual Women in a Male World*, 2006.
- [3] J. Carpenter, J. M. Davis, N. Erwin-Stewart, T. R. Lee, J. D. Bransford, and N. Vye, "Gender representation and humanoid robots designed for domestic use," *Int. J. Soc. Robot.*, vol.1, 2009, pp. 261-265.
- [4] P. Schermerhorn, M. Scheutz, and C. R. Crowell, "Robot social presence and gender: Do females view robots differently than males?" in *5th ACM/IEEE Int. Conf. Human-Robot Interaction*, 2008, pp. 263-270.
- [5] A. Powers, A. D. I. Kramer, S. Lim, J. Kuo, S. Lee, and S. Kiesler, "Eliciting information from people with a gendered humanoid robot," in *14th IEEE Int. Ws. Robot Hum. Interact. Commun. (RO-MAN)*, 2005, pp. 158-163.
- [6] C. Crowell, M. Scheutz, P. Schermerhorn, and M. Villano, "Gendered voice and robot entities: Perceptions and reactions of male and female subjects," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2009, pp. 3735-3741.

- [7] M. C. Siegel, C. Breazeal, and M.I. Norton, "Persuasive Robotics: The influence of robot gender on human behavior," IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), 2009, pp.2563-2568.
- [8] J. Rovertoson, "Gendering Humanoid Robots: Robo-Sexism in Japan," *Body Society*, vol. 16, no. 2, 2010, pp.1-36.
- [9] J. Weber, "Helpless machines and true loving care givers: A feminist critique of recent trends in human-robot interaction," *Info, Comm & Ethics in Society*, vol.3, 2005, pp.209-218.
- [1 0] J. Weber and C. Bath, "'Social' Robots & 'Emotional' Software Agents: Gendering Processes and De-Gendering Strategies for 'Technologies in the Making'," *Gender Designs IT: Construction and Deconstruction of Information Society Technology* (I. Zorn, S. Maass, E. Rommes, C. Schirmer, H. Schelhowe, eds), 2007, pp.53-63.
- [1 1] T. Nomura and S. Takagi, "Exploring Effects of Educational Backgrounds and Gender in Human-Robot Interaction," in *Int. Conf. User Science and Engineering (iUser)*, 2011, pp.24-29.
- [1 2] M. Kanoh, Y. Oida, Y. Nomura, A. Araki, Y. Konagaya, K. Ihara, T. Shimizu and K. Kimura, "Examination of Practicability of Robot Assisted Activity Program using Communication Robot for Elderly People," *J. Robotics and Mechatronics*, 2011, vol.23, no.1, pp.3-12.
- [1 3] Y. Ito, "Evaluation of Sex Roles as a Function of Sex and Role Expectation," *Jap. J. Educ. Psychol.*, 1978, vol.26, no.1, pp.1-11. (in Japanese).