

Evaluating the Nao Robot in the Role of Personal Assistant: The Effect of Gender in Robot Performance Evaluation [†]

Adrian Vega ^{*}, Kryscia Ramírez-Benavides, Luis A. Guerrero and Gustavo López

Centro de Investigaciones en Tecnologías de la Información y Comunicación, Universidad de Costa Rica, 2060 San José, Costa Rica; kryscia.ramirez@ucr.ac.cr (K.R.-B.); luis.guerrero@ecci.ucr.ac.cr (L.A.G.); gustavo.lopez_h@ucr.ac.cr (G.L.)

^{*} Correspondence: adrian.vegavega@ucr.ac.cr

[†] Presented at the 13th International Conference on Ubiquitous Computing and Ambient Intelligence UCAmI 2019, Toledo, Spain, 2–5 December 2019.

Published: 20 November 2019

Abstract: By using techniques such as the Wizard of Oz (WoZ) and video capture, this paper evaluated the performance of the Nao Robot in the role of a personal assistant, which was valued alongside the impact of the assigned gender (male/female) in the perceived performance of the robot assistant. Within a sample size of 39 computer sciences students, this study assessed criteria such as: perceived enjoyment, intention to use, perceived sociability, trust, intelligence, animacy, anthropomorphism, and sympathy, utilizing testing tools such as Unified Theory of Acceptance and Use of Technology (UTAUT) and Godspeed Questionnaire (GSQ). These methods identified a significant effect of the gender assigned to the robot in variables such as intelligence and sympathy.

Keywords: human–robot interaction; HRI; HCI; UTAUT; GSQ; WoZ; gender

1. Introduction

Starting as manufacturing machines, robots have opened to a wider application of areas such as education [1–3], personal assistance [4], health [5,6], and many others. Multiple research studies have described different elements to consider in the design of interactions with socially intelligent robots [6–9]. The theory of “Social Intelligent Robots” [6] establishes that the interaction with a robot must meet four different criteria. Being Socially Evocative relies on anthropomorphizing and capitalizes on feelings evoked. Being Socially Situated requires being able to react to other social agents and objects in the environment. The ability to be sociable and Proactively Engage with humans satisfies internal social aims. Displaying Social Intelligence requires showing deep models of human cognition and social competence during robot interaction.

One significant social element in the Human-Robot Interaction (HRI) is the gender assigned to the robot. Multiple research studies [10–16] have demonstrated that stereotypes transferred to the robot through the gender played a significant role in the user’s interaction, showing how the gender has an effect on the user perception of the robot’s performance and attitude executing tasks stereotypically associated to each gender [10]. Similar tests were also applied to different robot models aiming to the same results, and also demonstrated that participants feel more comfortable interacting with robots aligned with the stereotypical gender roles [11]. From this general postulation, gender has an effect in the user–robot interaction, and other research approaches had developed novel methods in the field, demonstrating the differences in the elicited information in the users when interacting with robots of different genders [12] or the level of persuasiveness assigned to the robot regarding the robot and user’s gender [13]. The effect of the robot gender can

be perceived even by changing only the isolated variables of voice levels [14] or aesthetic appeal [15] in the robot.

The current research evaluated the effectiveness of the robot Nao in the execution of the role of personal assistant, using criterias such as: trust, intention to use, perceived enjoyment, and perceived sociability. In addition, this work evaluated the impact of the gender in the perception of this performance by measuring the perceived enjoyment, intention to use, perceived sociability, trust, intelligence, animacy, anthropomorphism, sympathy, service value, and topic comforts.

2. Methods and Materials

2.1. Wizard of Oz (WoZ)

The Woz is utilized in technologies especially under development. In HRI, the WoZ technique [17] involves simulating the end state of the technology operating through the assistance of people operating the robot. This simulation creates the illusion that the robot works autonomously.

2.2. Video-Based Human–Robot Interaction (VHRI)

The VHRI [18] technique is one of the easiest and affordable techniques, and also allows the evaluation of the interaction with advanced technologies in controlled environments. The technique is used by recording videos of the desired interaction with the robot.

2.3. Unified Theory of Acceptance and Use of Technology (UTAUT)

UTAUT [19] is a theoretical framework that was created to evaluate the acceptance and usage of technologies. It integrates elements of other theories such as the theory of reasoned action, motivation model, theory of social cognition, and innovation diffusion theory, among others [15].

2.4. Godspeed Questionnaire (GSQ)

The GSQ [20] is a standardized instrument in the HRI field translated to multiple languages. It is supported by multicultural previous research. By using semiotic differential scales, the GSQ evaluates constructs such as anthropomorphism, animacy, likeability, and perceived intelligence.

2.5. The Robot

For this research, we used the Robot Nao Model V6 (see Figure 1) from SoftBank Robotic [21]. The main characteristics of this robot are: 58 cm height, 25 degrees of freedom in limb movement, 4 directional speakers, a limited manipulation of objects, verbal communication capacity, and humanoid appeal.



Figure 1. Nao robot picture.

3. Design

For this research, we designed two different scenarios:

3.1. First Scenario: Gender Evaluation Effect

In the first scenario, we applied the video capture technique and the GSQ, evaluating the impact of the gender assigned to the robot in its perceived performance as a personal assistant. For this scenario, we created two sets of videos of the robot performing the same activities, dialogues, and movements. The only difference was the voice tone of the robot in each video. In one pack of videos the voice emulated a male tone, while the other used a female appeal. The participants group was randomly divided into two cohort subgroups. Using headphones, each subgroup watched a single gender pack of videos (male or female). During the set of videos, the robot performed the same activities related to the role of personal assistant such as taking notes in a meeting or tracking agenda activities. After watching the videos, the participant completed a survey. In the survey, the robot was addressed consistently to the specific gender of the videos presented.

The survey used in this first part included items of the GSQ, plus some additional items related to the gender effect evaluation. Among the evaluated items are: animacy, anthropomorphism, intelligence, sympathy, value of the service, robot role, and type of information shared.

3.2. Second Scenario: Interaction Evaluation

In the second scenario, we applied the WoZ technique and the UTAUT questionnaire items. In this scenario, the participants interacted directly with the robot in a building reception area. In this scenario, the robot performed as a personal assistant. The robot provided information regarding the university courses, study programs, and professors' office locations, among other administrative information. The flow of this activity was related to the interaction and questions of the participants. No gender variables were evaluated in this scenario, and the robot presented with a standard androgynous voice.

For the scenario, the operator of the robot had a script for the most common questions and answers. In case the participants had other specific questions, we also included the capability of inputting a specific answer to the robot through typing unique sentences.

After the interaction, the participants completed a second survey evaluating their second experience with the robot. For this evaluation, we applied criterias of the UTAUT questionnaire such as intention to use, perceived enjoyment, perceived sociability, and trust.

3.3. Population

Both scenarios were applied in a group of 39 students of computer sciences from two different groups, which were divided in 30 male and nine female participants. The mean age was 21.21 years old with $SD = 2.32$. From the total of 39 participants, 17 participants received the video of Nao as a female personal assistant, while the other 22 participants received the video of Nao as a male personal assistant.

4. Results

4.1. First Scenario: Gender Evaluation Effect

According to the multiple one-way ANOVA analysis applied to the group, significant variance was identified in the analysis of sympathy $F(1,37) = 6.13$, $p = 0.018$ and intelligence $F(1,37) = 4.47$, $p = 0.036$, $\alpha = 0.05$. In both categories, the robot in the female role achieved higher mean values in the evaluation.

No significant differences were identified in the perception of animacy $F(1,37) = 2.30$, $p = 0.138$ and anthropomorphism $F(1,37) = 3.33$, $p = 0.76$, $\alpha = 0.05$.

Regarding the significant variance identified in the sympathy perception, higher values were achieved by the female robot, reaching $M = 4.39$, $SD = 0.65$; meanwhile, the male robot obtained $M = 3.76$, $SD = 0.87$. Analyzing the items in the sympathy category, the higher scores for the female robot were achieved for the items related to being friendly, kind, and pleasant.

Regarding the intelligence perceived, the higher values were also reached by the female robot succeeding $M = 3.98$, $SD = 0.77$; meanwhile, the male robot reached values of $M = 3.34$, $SD = 0.87$. The

items for which the female robot highly succeeded in the intelligence category were related to an evaluation of responsibility and being reasonable.

No significant difference was identified in the evaluation of the value of the service provided according to the gender of the robot $F(1,37) = 0.33, p = 0.571, \alpha = 0.05$. The female service value perception was $M = 623.47, SD = 314.30$, while the male service value perception was $M = 270.45, SD = 294.66$ (see Figure 2).

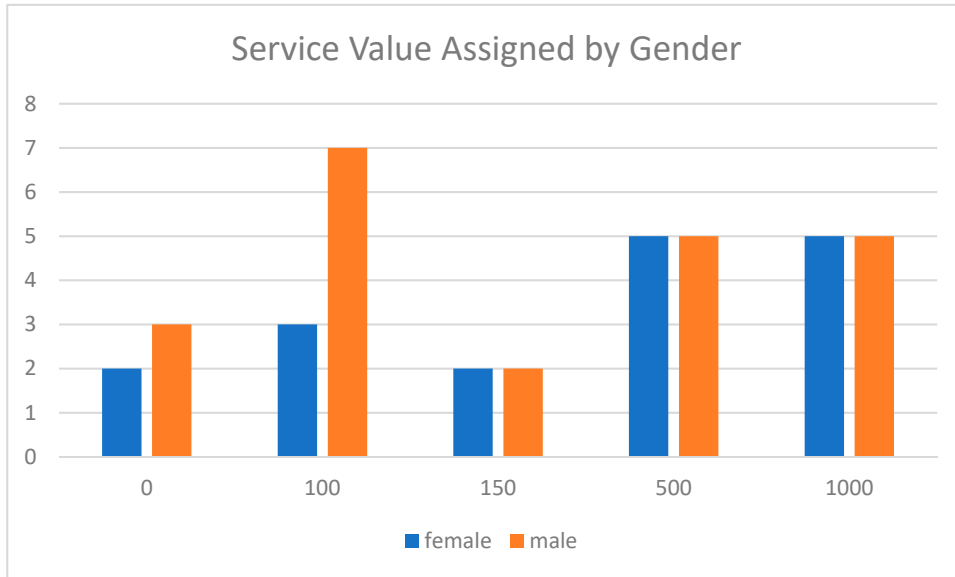


Figure 2. Number of participants x value assigned to the service, separated by gender. Currency = Costa Rican colones.

Finally, no significant differences were identified regarding the kind of information the participants felt comfortable sharing with the robot related to the gender (see Figure 3). However, there was a significant difference in the information that the participants felt comfortable sharing; “data and information” was the most common material that the participants were willing to share with the robot.

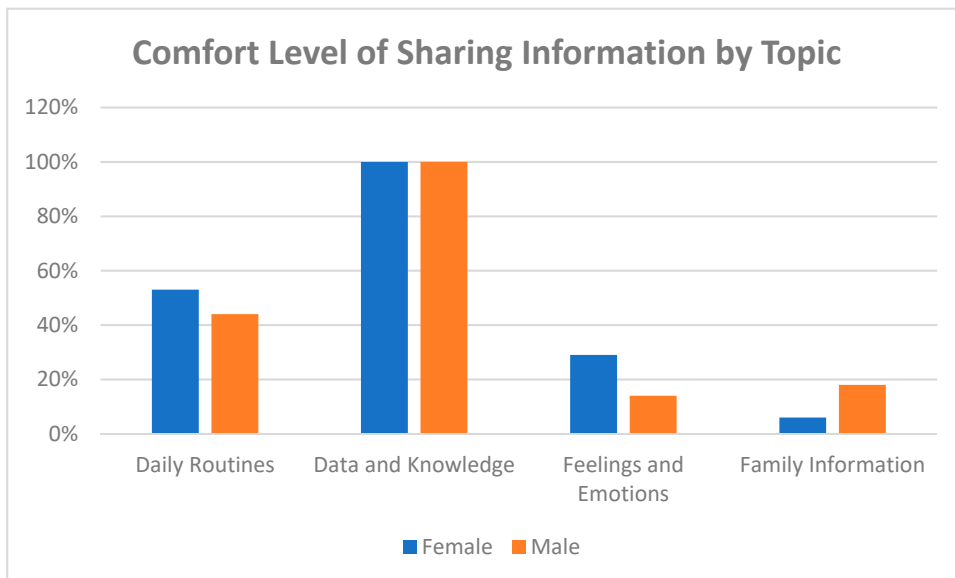


Figure 3. Level of comfort that participants felt sharing information with the robot, sorted by gender and type.

4.2. Second Scenario: Robot Performance Perception

By using a Likert scale with 5 levels, in which 1 is the least positive, 5 is the most positive, and 3 neutral. The participants evaluated mostly positively the interaction with the robot when it was performing the role of an assistant. The higher perception related to the intention to use the robot in other activities ($M = 4.41$, $SD = 0.72$). Similar values was achieved regarding the perception of enjoyment ($M = 4.27$, $SD = 0.72$) and perceived sociability ($M = 4.20$, $SD = 0.65$). A less satisfactory value was related to the perception of trust in the robot during the developed activity ($M = 3.86$, $SD = 0.93$) (see Table 1).

Table 1. Evaluation of the direct interaction with the Nao robot through Wizard of Oz (WoZ).

Variable	N	Mean	S.E. Mean	Std Dev	Kurtosis	S.E. Kurt	Skewness	S.E. Skew	Minimum
IntentionUse	39	4.41	0.12	0.72	2.40	0.74	-1.59	0.38	2.33
PerceivedEnjoyment	39	4.27	0.10	0.65	-0.80	0.74	-0.56	0.38	2.80
PerceivedSociability	39	4.20	0.10	0.65	-0.31	0.74	-0.71	0.38	2.75
Trust	39	3.86	0.15	0.93	-0.50	0.74	-0.52	0.38	2.00

5. Conclusions

The performance of the robot as an assistant was evaluated during the direct interaction of the WoZ. It performed with significant highly performance values. The levels of sociability, enjoyment, intention to use, and trust were satisfactory according to the standards of the test. It might be said that the level of trust is continuously the less favorable category according to the UTAUT standards. The lack of trust in the robot or technology is a challenge to develop in this kind of interaction. It should also consider the evaluation of the exposition of this kind of technology regarding time as a variable.

Regarding the gender role evaluated, we identified significant differences in the intelligence and sympathy perceived. The robot representing the female role achieved higher values. No significant differences were identified in the animacy, anthropomorphism, service value, and topic comfort.

The higher intelligence variability ratings for the female robot were related to the perception of responsibility and reasonability in the female assistant. Regarding sympathy, the female assistant was perceived as being more kind, friendly, and pleasant.

6. Analysis

This research evaluated the effectiveness of the Nao robot in the role of personal assistant. The results are consistent with the previous references. The decrease of trust, in relation to the other variables such as intention to use, is consistent with the lack of will to sharing meaningful information with robots. It was most preferable to share only information related to “information and data”. It appears as though the social robots are still struggling to obtain trustworthiness from participants over functionality. Although their performance during tasks is impressive, more solid reliance at deeper levels is still required.

Regarding the significant effect that the robot’s gender has the human interaction, this study shows a significant difference in the perception of intelligence and sympathy. Further investigation is required to elaborate the effect of the selected role, which is conventionally assigned to the female role, and the mental models linked to the gender itself. According to the current results, during the female representation of the role, the attribution of intelligence could be related to the higher perception of responsibility and reasonability; meanwhile, the higher sympathy rating might be related to a higher perception of being friendly, kind, and pleasant. This is consistent with the previous research stating that records of better performance are assigned to robots executing tasks aligned to traditional roles. However, more research is required to state if that is the only factor causing this significant difference, or if there is a more positive general perception of the female gender regarding the intelligence and sympathy variables.

Regarding the non-significant differences identified in the service value, it is also interesting to mention that no significant gap was identified between genders, even though one of them was better qualified in its performance. Similar research might be also important to identify gaps in the gender perception in different roles and scenarios.

Finally, no significant differences were found in the variables of animacy and anthropomorphism, which might be explained because no changes were made in the aesthetics of the robot or movements. The reliability of the test was consistently maintained in both videos, when no significant difference was identified related to the gender perception.

As a risk internal evaluation of the current experiment, it is important to consider that the population surveyed was only 39 people. Most of them were male, and all were students of computer science. It is required that future research improve the tests applied in this experiment and complement the results discovered. In addition, enlarging the scenario with a wider spectrum of roles might also improve the ecological validity of the results.

Acknowledgments: This work was supported by the CITIC-UCR (Centro de Investigaciones en Tecnologías de la Información y Comunicación) and by ECCI-UCR (Escuela de Ciencias de la Computación e Informática), grand No. 834-B7-267. Thanks to the User Interaction Group (USING) for supporting the research.

References

1. Ramírez-Benavides, K.; López, G.; Guerrero, L.A. *Designing Tools that Allows Children in the Early Childhood to Program Robots*; Springer: Cham, Switzerland, 2017; pp. 71–89.
2. Bravo, F.A.; González, A.M.; González, E. Interactive Drama with Robots for Teaching Non-Technical Subjects. *J. Human-Robot Interact.* **2017**, *6*, 48–69.
3. Baxter, P.; Ashurst, E.; Read, R.; Kennedy, J.; Belpaeme, T. Robot education peers in a situated primary school study: Personalisation promotes child learning. *PLoS ONE* **2017**, *12*, 1–23.
4. Torta, E.; Oberzaucher, J.; Werner, F.; Cuijpers, R.H.; Juola, J.F. Attitudes Towards Socially Assistive Robots in Intelligent Homes: Results from Laboratory Studies and Field Trials. *J. Human-Robot Interact.* **2013**, *1*, 76–99.
5. Wood, L.J.; Dautenhahn, K.; Rainer, A.; Robins, B.; Lehmann, H.; Syrdal, D.S. Robot-Mediated Interviews—How Effective Is a Humanoid Robot as a Tool for Interviewing Young Children? *PLoS ONE* **2013**, *8*, e59448.
6. Dautenhahn, K. Socially intelligent robots: Dimensions of human-robot interaction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2007**, *362*, 679–704.
7. Fortunati, L. *Social Robots from a Human Perspective*; Springer: Berlin, Germany, 2015; Volume 73, pp. 1–144.
8. Dautenhahn, K. Roles and functions of robots in human society: Implications from research in autism therapy. *Robotica* **2003**, *21*, 443–452.
9. Vlachos, E.; Jochum, E.; Demers, L.-P. The effects of exposure to different social robots on attitudes toward preferences. *Interact. Stud.* **2016**, *17*, 390–404.
10. Kuchenbrandt, D.; Häring, M.; Eichberg, J.; Eyssel, F.; André, E. Keep an Eye on the Task! How Gender Typicality of Tasks Influence Human-Robot Interactions. *Int. J. Soc. Robot.* **2014**, *6*, 417–427.
11. Tay, B.; Jung, Y.; Park, T. When stereotypes meet robots: The double-edge sword of robot gender and personality in human-robot interaction. *Comput. Human Behav.* **2014**, *38*, 75–84.
12. Powers, A.; Kramer, A.D.I.; Lim, S.; Kuo, J.; Lee, S.L.; Kiesler, S. Eliciting information from people with a gendered humanoid robot. *Proc.—IEEE Int. Work. Robot Hum. Interact. Commun.* **2005**, *2005*, 158–163.
13. Siegel, M.; Breazeal, C.; Norton, M.I. Persuasive robotics: The influence of robot gender on human behavior. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009, St. Louis, MO, USA, 11–15 October 2009; pp. 2563–2568.
14. Nass, C.; Moon, Y.; Green, N. Are machines gender neutral? Gender-stereotypic responses to computers with voices. *J. Appl. Soc. Psychol.* **1997**, *27*, 864–876.
15. Eyssel, F.; Hegel, F. (S)he’s Got the Look: Gender Stereotyping of Robots. *J. Appl. Soc. Psychol.* **2012**, *42*, 2213–2230.
16. Jung, E.H.; Waddell, T.F.; Sundar, S.S. Feminizing Robots. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA ’16, San Jose, CA, USA, 7–12 May 2016; pp. 3107–3113.

17. Usability Body of Knowledge, Wizard of Oz, 2012. Available online: <http://www.usabilitybok.org/wizard-of-oz> (accessed on 13 April 2017).
18. Walters, M.L.; Lohse, M.; Hanheide, M.; Wrede, B.; Syrdal, D.S.; Koay, K.L.; Green, A.; Huttenrauch, H.; Dautenhahn, K.; Sagerer, G.; et al. Evaluating the Robot Personality and Verbal Behavior of Domestic Robots Using Video-Based Studies. *Househ. Serv. Robot.* **2014**, *25*, 467–486.
19. Oshlyansky, L.; Cairns, P.; Thimbleby, H. Validating the Unified Theory of Acceptance and Use of Technology (UTAUT) Tool Cross-Culturally. In Proceedings of the 21st British HCI Group Annual Conference on HCI 2007: HCI...but not as We Know It—Volume 2, Lancaster, UK, 3–7 September 2007.
20. 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.* **2009**, *1*, 71–81.
21. Robotics, S. Nao, 2019. Available online: <https://www.softbankrobotics.com/emea/en/nao> (accessed on 19 September 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).