2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Smart Speaker vs. Social Robot in a Case of Hotel Room

Junya Nakanishi[1], Jun Baba[2], Itaru Kuramoto[3], Kohei Ogawa[4], Yuichiro Yoshikawa[1] and Hroshi Ishiguro[1]

***Abstract*— Under the circumstances that social robots are increasingly being developed and studied in service encounters at public spaces, are they introduced into residential environments (i.e., private space)? This study hypothesizes that a personal assistant device in residential environments should wear human-like appearance to engage in service as conversation partner. We implemented the interaction design that provides regular services as the current personal assistant and additional service as conversation partner, and then conducted a field experiment where the participants stayed in the hotel rooms with a smart speaker or a social robot. The results support the hypothesis in that of conversation amount and emotional experience by conversation. The results also suggest the possibility of commercial service, namely conversational advertisement through social robots.**

I. INTRODUCTION

In service encounters at public spaces, social robots are increasingly being developed and studied. A social robot is a physically embodied agent that interact and communicate with humans by social behavior using humanoid parts (e.g., torso, head, face, or arms). Researchers developed such social robots for service as museum guides [1], travel guides [2], salespersons [3], and receptionists [4]. Service industry expects such robots to substitute human labor or extend service value [5], [6].

As a conversational service robot at residential environments (i.e., private space), smart speakers (e.g., Google Home and Amazon Echo) are increasingly being used [7], [8]. A smart speaker is a voice-controlled speaker with an integrated personal assistant that offers a wide range of service such as information seeking, playing music and chat. Most existing smart speakers have simple appearance such as cube and column that seems to be machine-like rather than human-like objects as a social robot. Here, a question arises: is such a personal assistant device wearing machine-like appearance better than human-like appearance in residential environments?

Studies of service robots' appearance found differences between machine-like and human-like robots in human-robot interaction. For example, human-like robots tend to be perceived as friendlier, more intelligent and autonomous than machine-like [9], [10], [11]. People tend to engage more in emotional and social conversation with human-like robots than machine-like [12], [13]. People tend to be more satisfied with the service provided by machine-like robots than human-like robots [12], [14]. Summarizing the above, human-like appearance may be required more when a personal assistant plays a role of conversational partner for users, not when it plays a role of a tool, which only engages in requested service. However, those studies were limited in a few single-time human-robot interaction or impression evaluation via videos or images. That is to say, we do not have enough knowledge to answer the difference in residential environments, where a same user interacts with the robot several times on a varied context. Repeated interaction and real context may refract the effects of machine-like or human-like appearance.

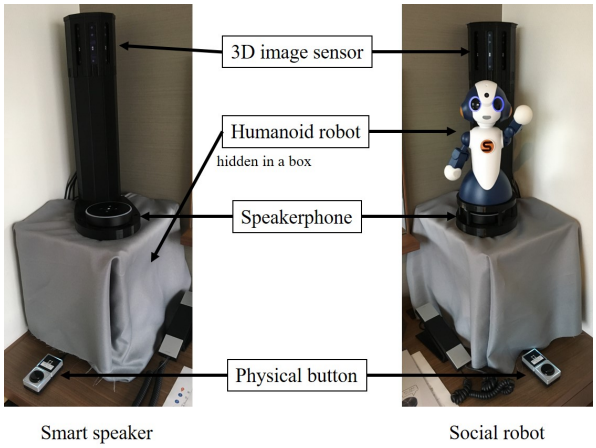

Fig. 1: Personal assistant devices

This study focuses a case of a hotel room as one of the residential environments. In a hotel context, a conversational service robot is potential service as conversation partner, such as short chatting and heartwarming interaction [15]. To investigate the necessity of human-likeness, we conducted a field experiment where a guest shared a hotel room with a smart speaker or a social robot for a night. A smart speaker and a social robot provide same verbal communication, which means that the only one difference is their appearance: a column and a humanoid object with physical behavior (Fig. 1). The main contribution of the paper is evaluation with longitudinal date (e.g., user behavior over time) in practical environment to verify a hypothesis: a personal assistant device in residential environments should wear human-like appearance to engage in service as conversation partner.

*This work was supported by Tokyu Resorts&Stay Co., Ltd, Tokyu Fudosan R&D Center Inc. and the Grant-in-Aid for research activity start-up, JSPS KAKENHI Grant Number 18H06479.

[1]Junya Nakanishi, Yuichiro Yoshikawa and Hiroshi Ishiguro are with Osaka University, Osaka 560–8531, Japan {nakanishi, yoshikawa, ishiguro}@irl.sys.es.osaka-u.ac.jp

[2]Jun Baba is with Cyber Agent Inc., Tokyo 150–0042, Japan baba_jun@cyberagent.co.jp

[3]Itaru Kuramoto is with The University of Fukuchiyama, Kyoto 620–0886, Japan kuramoto-itaru@fukuchiyama.ac.jp

[4]Kohei Ogawa is with Nagoya University, Aichi 464–8603, Japan k.ogawa@gmail.com

II. Personal Assistant in Hotel Room

A. Hardware

The hardware system consists of three-dimensional (3D) image sensors, a microphone array with a speaker (speakerphone), a physical button, and a humanoid robot (Fig. 1). For the 3D image sensor, we used RealSense D415 (Intel Corp.), which can provide color (RGB) and depth images (maximum range of more than 10 m). For the speakerphone, we used M1 (eMeet), which can recognize voices (maximum range of more than 5 m) with an acoustic echo cancellation feature. For the physical button, we used a button included Palatte Kit (Grant & Union Inc.), which allows a user to physically input a signal. For the humanoid robot, we used Sota (Vstone Co., Ltd.), which is a desktop-sized (about 0.3 m tall) robot capable of interacting with a human using its voice, arms, face, and head motion. In case of smart speaker condition, the humanoid robot is hidden in a box under the displayed system, and its physical behavior is stopped. The other instruments (e.g., a control computer, power socket, and a Wi-Fi rooter) are also hidden in the box.

The reason why the humanoid robot is employed as a representative of human-like robots is that it is medium-level human-likeness, which means here that it is not realistically human-like but possess some human-like features in appearance. The best known hypothesis of human-likeness, Uncanny Valley hypothesis [16], says that high level of human-likeness is difficult to avoid uncanny feeling. The uncanny feeling seems to be not suitable for the service. In fact, impression surveys suggested that users do not hope realistically human-likeness to robot's appearance for domestic applications [9]. Besides, past studies of social robots' appearance investigated differences between low and medium-level human-likeness, which are called as "mechanoid and humanoid robots" or "product-ness and robot-ness" [9], [12]. Thus, this study focuses the appearance differences between speakerphone and humanoid robot in residential environments.

B. Software

The internal system are based on updated previous system [15]. A scenario manager is the central software that deals with the human-robot interaction. It continuously receives the situational data from the sensors, i.e., the 3D image sensor, the speakerphone, the physical button, and the humanoid robot, and sends behavioral commands to the humanoid robot based on the scenario. When the scenario manager receives data in accordance with the present interaction state, a state transition occurs. Then, the scenario manager picks up the appropriate scenario from the scenario database by referring to a state transition table, and sends the behavioral commands to the robot. The behavioral commands include utterances, physical motions (e.g., bowing, raising a hand, or nodding), changing the color of the eye (e.g., red, blue, or green), and tracking human face by directing its own. The robot performs these behavioral commands using a voice synthesis or a motion database. Note that the robot's voice is played through the speakerphone.

The 3D image sensor system sends the 3D position of the customer and brightness of the room to the scenario manager. From the captured RGB image, it detects the two-dimensional (2D) position of the user using the OpenPose algorithm [17], and computes the human 3D positions with the detected 2D positions and the depth image. Using the 3D position, we judged guest's behavioral situation: front of a desk, sleeping in bed, staying up in bed, and passage. Brightness is computed with the RGB image, and is used for judging whether the lights are on or off. The speakerphone system sends the recognized voice keywords to the scenario manager. Voice recognition software Cloud Speech-to-Text (Google LLC), converts the recorded voice into text. Keywords related to the scenario are extracted from the text and subsequently sent to the scenario manager. The physical button system tells guest's pressing it to the scenario manager. The button controls permission of the system's proactive utterance.

C. Interaction Design

Aming at constructing the interaction design that provides regular services as the current personal assistant and additional service as conversation partner, we implemented the hotel room's personal assistant system which engages in short chatting with guests, support of information seeking (i.e., question answering and recommendation), and heartwarming interaction. Support of information seeking is regular services as the current personal assistant [7], [8]. Chatting and heartwarming interaction will be services as conversation partner because they demand a partner of such personal things as character, preference, and thoughtfulness unlike just a tool.

1) Basic Interaction: The assistant mainly has two state: wait and active state. It is generally in wait state; namely, waiting for events (e.g., user's calling and environmental change). In case of the social robot, it faces to a wall to avoid evoking user's feeling of be watched by it while wait state. If a certain event happens, the assistance makes utterance and changes wait state into active state. After making utterance, the assistant waits for 20 seconds until a user makes response. When 20 seconds pass, it makes additional proactive utterance in accordance with the context (see hereafter section II-C.4 for detail) or wait for more 15 seconds and then says "*Please call me again*" while going into wait state. Note that changing wait state into active state without user's calling can be prohibited by the physical button. In other words, the button can order the assistant to refrain from starting conversation unless a user calls it.

The assistant makes proactive and passive utterance. Proactive utterance is triggered by guest's motion, elapsed time, and environment change, such as guest entering, several seconds after certain conversation ends, and light off. With the proactive utterance, users feel like being talked by the assistant. Proactive utterance will be important for assistant to be conversation partner for users. In this system, all proactive utterance belongs to heartwarming interaction.

Passive utterance is triggered by user's utterance. The assistant make a response to a detected keyword in user's utterance. When not preparing a corresponding function, such as "*Please turn off the light*" and "*Please turn on the television*", the assistant says "*Sorry, it will be implemented in near future.*" When not preparing a corresponding response, the assistant says "*Sorry, I could not understand what you say.*" Note that the responding function does not work in wait state.

To go into active state from wait state, there are two ways: waiting the assistance's proactive utterance and calling the assistant's name Sota (e.g., "*Hey, Sota*"). The word is called the magic words or the wake words. When only name is called, it responds by saying "*Yes! May I help you?*, and goes into active state. In the case of the social robot, it starts tracking the guest's face at the same time.

In addition to the above, this system has some more functions; at the first meet, the assistant explains how it can be interacted with; it turns off power from 2 to 6 a.m. or when the user is asleep; it tells the user the c state of proactive utterance when the user operates the button; and finally, it demonstrates it is active by eye light, showing breathing motions, and movement of arms and neck during interaction with user.

2) Short Chatting: To be a communication partner for users, this assistant engages in short chatting: a simple conversation, self-introduction, dance performance, a simple game, and trivia introduction.

For A simple conversation includes replies to the greetings and the farewells (e.g., "*Good morning*", "*Good night*", and "*How are you?*"), replies to thanking (e.g., "*You are welcome*" and "*Not at all*"), and other replies (e.g., "*Good job today. Please relax in this hotel*" in reply to "*I'm tired*"). For self-introduction, the assistant is capable of answering a simple question about itself, such as age, hobby, gender, and birthplace.

When a guest says "I'm bored.", the assistant randomly proposes playing together, namely showing dance performance, a simple game, and trivia introduction. For dance performance, the assistant performs a simple dance while singing. Afterwards, it asks a user whether good or bad the dance is. For a simple game, the assistant starts a rock paper scissors game. For trivia introduction, the assistant talks randomly about one of four trivia, and answers to a few questions about the trivia. It can also move on the next trivia introduction by "*Anything else?*"

3) Support of Information Seeking: Hotel guests want information about accommodation and local area [18]. Thus, the system answers questions about accommodation (22 items; e.g., breakfast, checkout, and card key), local area (17 items; e.g., local restaurants, souvenir shops, and tourist spots), and public information (11 items; e.g., weather, news, current time). As for local area, when requested, the assistant recommends one of prepared items (i.e., five restaurants, three souvenir shops, and three tourist spots). After recommendation, it can answer such the details as place, opening hours, and price. It can also move on the next recommendation by "*The other recommendations?*"

4) Heartwarming Interaction: Heartwarming interaction is the behavior and attitude that creates a heartwarming feeling, referred to as interpersonal warmth [19], [20]. Basically, that is proactive behavior relating to such intent as friendliness, helpfulness, sincerity, trustworthiness, and morality [21]. In service encounters, heartwarming interaction is important because it enhances customer satisfaction regarding the whole service experience and increases the number of frequent customers and positive word of mouth [22], [23].

To provide heartwarming feeling, the assistant proactively engages in greetings, farewells, and warm comments as a previous study [18]. The assistant says greetings and farewells as per the guest's state the 3D image sensor judged (see section II-B for detail). When the guest enters the room, the assistant says "*Welcome back*"; when the lights are turned off, the assistant says "*Good night*"; and when the guest wakes up in the morning, the assistant says "*Good morning*" The assistant sometimes says warm comments involving cautionary note in stay, accommodation in a hotel, local information (e.g., restaurant, traffic, news, and weather). They were provided after greetings and farewells, silent 20 seconds after the assistant's utterance ends, or after the assistant apologizes for not understanding guest's utterance. For example, the assistant says "*Welcome back! Is the room cold? The remote control for the air conditioner is near the TV, so please adjust it if you want*", "(silent 20 seconds) *Well, would you like a fruit cocktail with seasonal fruits? If you are interested, please let me know. I can introduce a nice bar*", or "*Sorry, I could not understand what you say. Instead, I will notify important information for you. Please be careful not to forget your room key when you go out. The room door is automatically locked.*"

III. Field Experiment

We conducted a field experiment where the participants stayed in the hotel rooms with a smart speaker or a social robot. The purpose of the experiment is gathering date in practical environment to verify a hypothesis: a personal assistant device in residential environments should wear human-like appearance to engage in service as conversation partner. We gathered guests' behavior data and the impressions of the personal assistant devices. The experiment was conducted at a business hotel in Tokyo for three weeks during the September of 2019. This study was approved by the ethics committee of the Graduate School of Engineering Science in Osaka University.

A. Methodology

We used three rooms in the fourth floor of the hotel. The layout of all the rooms is almost the same, although bilaterally asymmetrical. They were single occupancy rooms, each of which was approximately 16 sq. m in area, furnished with a double-sized bed, desk, television set, air conditioner, air cleaner, microwave oven, and bath room. The personal assistant devices were placed on a desk.

TABLE I: Items in the questionnaire.

Item	*Statement*
W	The assistant was [X]. [X]:= friendly, trustworthy, kind, sincere, and warm
C	The assistant was [Y]. [Y]:= skillful, intelligent, proving enough information, and competent
CP1	The conversation with the assistant was fun.
CP2	I did not feel lonely because of the assistant.
P1	I felt that the assistant was invading my privacy.
P2	I felt that it was difficult to take private behavior because of the assistant.

We recruited 60 Japanese participants (all aged between 20–60) through temporary staffing agency. The half (12 males and 18 females) were assigned to a smart speaker condition, the others (18 males and 12 females) were assigned to a social robot condition. We explained the purpose that we evaluate the personal assistant device in a hotel room by showing the actual devices. We also provided assurance on the protection of their private data and explained that their images and audio data would be saved and used for analysis. After receiving the informed consents from the participants, we explained how to interact with the assistant for detail. We instructed the participants to use the hotel as they would normally use and cooperate on a questionnaire and an interview at their checkout. Participants checked in the hotel from 5 p.m. to 8 p.m., stayed for a night, and checked out from 7 a.m. to 9 a.m.

B. Measurement and Analysis

If the personal assistant is regarded as a conversation partner, amount of short chatting will increase. As for heartwarming interaction, participants will feel more interpersonal warmth from the assistant. We measured them and compared with Wilcoxon rank sum test. Incidentally, the related items were also checked as well.

1) Records of the assistant's behavior: We analyzed the records of the assistant's behavior during the experiments. The records included the time, as well as the contents, of the assistant's responses. We counted the amount of assistant's utterance in every one hour, short chatting, information seeking conversation, heartwarming interaction.

2) Records of the sensors: We analyzed the records of the participants' position, utterance, and button operation during the experiments. Firstly, where a participant had conversation with the assistant was counted: front of a desk, in bed, and passage. Conversation is here defined as a series of utterances at intervals of not more than 40 seconds and including both participants' and the assistant's because the assistant goes into wait state 35 seconds after the last its utterance ends (see section II-C.1 for detail). Participant's conversation position was computed with the average of participant's position during the conversation. Besides, conversation was classified into two types: assistant-triggered and user-triggered conversation. Assistant-triggered conversation is defined as conversation that the assistant starts by proactive utterance. User-triggered conversation is defined as conversation that a user starts by a magic word. Secondly, we measured the time span when participants prohibited the assistant's proactive utterance by the records of button operation.

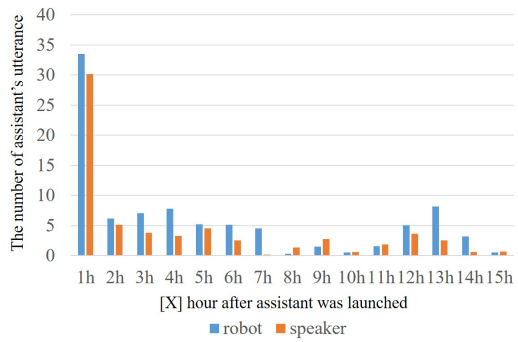

Fig. 2: The assistant's utterance frequency in every an hour

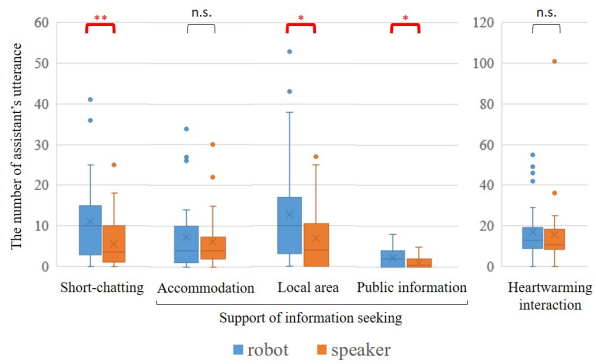

Fig. 3: Kinds and frequency of the assistant's utterance (*:p<.05, **:p<.01, hereafter)

3) Questionnaire: We collected the participants' impression of and user experience of stay with the personal assistant device using a questionnaire. The main questionnaire items are displayed in Table I. It enquires the participants about warmth (W), competence (C), being conversation partner (CP), and privacy invasion (P). Warmth and competence were measured by warmth-related and competence-related trait adjectives in previous studies [21], [24]. As for being conversation partner and privacy invasion, our associated items were used. A 7-point Likert scale was used to evaluate the questionnaire items (1: strongly disagree; 2: moderately disagree; 3: disagree a little 4; neither agree nor disagree, 5: agree a little; 6: moderately agree; 7: strongly agree).

C. Results

Figure 2 shows records of the assistant's utterance frequency history for participants' stay. The time scale shows elapsed time after assistant was launched by participants. The bar graph shows the participants' average number of assistant's utterance in each an hour. Conversation with the assistant concentrated on the first hour, and afterwards suddenly freed. In the time interval denoted by 7–11h, there were few conversation. Participants seem to sleep for those hours. The time interval denoted by 12–13h had an increase to some extent, whose utterance seems to be done after the participants waking up or before leaving.

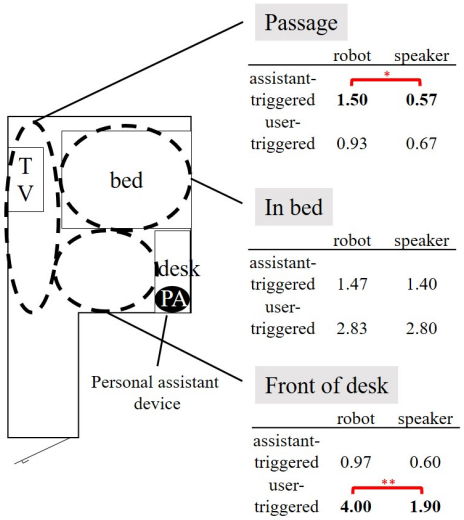

Fig. 4: Participant's conversation position and frequency

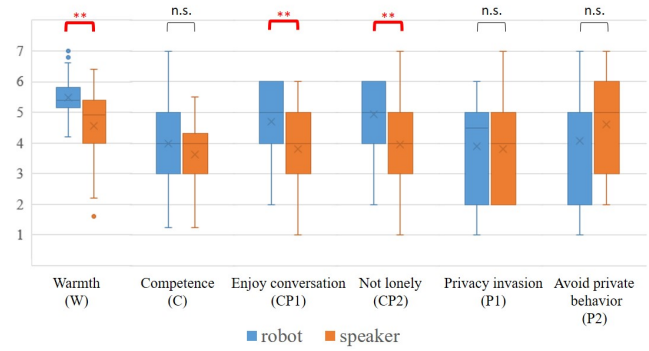

Fig. 5: Impression of interaction with the assistant

Figure 3 shows what kind of, as well as how much, the assistant's utterance took place for participants' stay. We found significant differences in short chatting (W=627, p=.0089), local area (W=595.5, p=.030), and public information (W=589, p=.033). The results suggest that participants and the assistant in robot condition had more conversation about them than speaker condition.

Figure 4 shows where participants talked with the assistant for participants' stay. We found significant differences in assistant-triggered conversation at the passage (W=595.5, p=.020) and user-triggered conversation at the front of a desk (W=636.5, p=.0053). The results suggest that participants in robot condition responded to the assistant's proactive utterance more than speaker condition at the passage and that participants in robot condition spoke to the assistant more than speaker condition in front of a desk.

Figure 5 shows participants' impression of interaction with the assistant. Regarding warmth and competence, the average score of items was used for analysis. Those scale was highly reliable with Cronbach's: warmth α = .82 for robot condition and .93 for speaker condition; competence α = .91 for a non-heartwarming condition and .86 for speaker condition. We found significant differences in warmth (W=675, p=.00085), fun of conversation (W=622.5, p=.0083), and not lonely (W=657.5, p=.0015). The results suggest that participants in robot condition felt more interpersonal warmth from the assistant than speaker condition and that participants in robot condition enjoyed conversation with the assistant more and felt less loneliness than speaker condition.

As for button operation, 4 participants prohibited the assistant's starting conversation unless they calls it for more 8 hours in robot condition; 15 participants did that in speaker condition.

IV. DISCUSSION

A. The necessity of human-likeness

Many results support the hypothesis: a personal assistant device in residential environments should wear human-like appearance to engage in service as conversation partner. First, human-likeness increases conversation chances between a guest and a personal assistant device. In fact, participants in this experiment spoke to the social robot more in front of it than the smart speaker; participants responded to the social robot more at the passage than the smart speaker. More conversation chances allow a personal assistant to pile a value of conversation service. The reason why human-likeness increases conversation chances would be that a human feels more affinity to humanoid devices than mechanical ones also in residential environments. In fact, participants in this experiment had more social conversation (i.e., short chatting) with the social robot than the smart speaker. The fact that the physical button was relatively used for the smart speaker suggests less affinity to mechanical devices than humanoid ones as a communication partner.

Second, human-likeness increases perceived interpersonal warmth from a personal assistant device. In fact, participants in this experiment felt more interpersonal warmth from the social robot than the smart speaker. As mentioned earlier, the emotional experience enhances customer satisfaction regarding the whole service experience and increases the number of frequent customers and positive word of mouth [22], [23]. As well, the other emotional experience in conversation might be increased by human-likeness. The reason why human-likeness increases perceived interpersonal warmth would be that a human imagines mind in human-like entities. In the other word, heartwarming interaction would be associated with friendly, helpful, or sincere intention of the assistant more by human-likeness.

Finally, participants enjoyed conversation with the social robot more and felt less loneliness than speaker condition. It is assumed that humanlike-ness enhances value of service as conversation partner, which seems to be caused by the above two factors: conversation amount and emotional experience by conversation. Therefore, we conclude that a personal assistant device in residential environments should wear human-like appearance to engage in service as conversation partner. Note that human-likeness is limited until medium level (i.e., humanoid) in our conclusion.

B. The possibility of commercial service

Toward commercial service, privacy concern regarding a humanoid robot is a potential problem in residential environments [15]. For example, a human may feel like that it observes his/her private behavior or that it steals his/her private information. In short, an increment of privacy concern may be a demerit of human-likeness in residential environments. However, the results show no significant difference in impression of privacy concerns. This would be achieved by the interaction design in which the social robot faces to a wall to avoid evoking user's feeling of be watched by it while wait state. Actually, a study suggests that turning the robot away from humans prevents the feeling of privacy invasion [25]. Likewise, using a technique of suppressing privacy concerns, social robots will be increasingly used more in residential environments.

Interestingly, frequency of asking the social robot about local area was more than the smart speaker. In travel, do you usually ask a receptionist to get local information? We conjecture that participants regarded the social robot as a local person because the robot seemed to "live" in the hotel. That is to say, human-like entity physically existing there might be suggestive of a person who is familiar with the local area. If so, this would be applied for advertisement. For example, if a social robot tells a blurb when requested to recommend, guests would accept it. Human-likeness contributes on an increment of advertising opportunity. The relationship between a guest and a robot might be important for an increment of its acceptance rate.

V. CONCLUSIONS

To verify a hypothesis: a personal assistant device in residential environments should wear human-like appearance to engage in service as conversation partner, we implemented the interaction design that provides regular services as the current personal assistant and additional service as conversation partner, and then conducted a field experiment where the participants stayed in the hotel rooms with a smart speaker or a social robot. The results support the hypothesis in that of conversation amount and emotional experience by conversation. The results also suggest the possibility of commercial service, namely conversational advertisement through social robots.

REFERENCES

[1] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Interactive humanoid robots for a science museum," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 2006, pp. 305–312.

[2] R. Triebel, K. Arras, R. Alami, L. Beyer, S. Breuers, R. Chatila, M. Chetouani, D. Cremers, V. Evers, M. Fiore, *et al.*, "Spencer: A socially aware service robot for passenger guidance and help in busy airports," in *Field and service robotics*. Springer, 2016, pp. 607–622.

[3] M. Watanabe, K. Ogawa, and H. Ishiguro, "Can androids be salespeople in the real world?" in *Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems*, 2015, pp. 781–788.

[4] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. C. Schultz, *et al.*, "Designing robots for long-term social interaction," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2005, pp. 1338–1343.

[5] S. H. Ivanov, C. Webster, and K. Berezina, "Adoption of robots and service automation by tourism and hospitality companies," 2017.

[6] H. Osawa, A. Ema, H. Hattori, N. Akiya, N. Kanzaki, A. Kubo, T. Koyama, and R. Ichise, "Analysis of robot hotel: Reconstruction of works with robots," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2017, pp. 219–223.

[7] A. Pyae and T. N. Joelsson, "Investigating the usability and user experiences of voice user interface: a case of google home smart speaker," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 2018, pp. 127–131.

[8] A. Berdasco, G. López, I. Diaz, L. Quesada, and L. A. Guerrero, "User experience comparison of intelligent personal assistants: Alexa, google assistant, siri and cortana," in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 31, no. 1, 2019, p. 51.

[9] M. L. Walters, K. L. Koay, D. S. Syrdal, K. Dautenhahn, and R. Te Boekhorst, "Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials," *Procs of New Frontiers in Human-Robot Interaction*, 2009.

[10] B. T. T. Chee, P. Taezoon, Q. Xu, J. Ng, and O. Tan, "Personality of social robots perceived through the appearance," *Work*, vol. 41, no. Supplement 1, pp. 272–276, 2012.

[11] M. Harbers, M. M. Peeters, and M. A. Neerincx, "Perceived autonomy of robots: effects of appearance and context," in *A World with Robots*. Springer, 2017, pp. 19–33.

[12] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, "All robots are not created equal: the design and perception of humanoid robot heads," in *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, 2002, pp. 321–326.

[13] C. Bartneck, T. Bleeker, J. Bun, P. Fens, and L. Riet, "The influence of robot anthropomorphism on the feelings of embarrassment when interacting with robots," *Paladyn, Journal of Behavioral Robotics*, vol. 1, no. 2, pp. 109–115, 2010.

[14] S. S. Kwak, "The impact of the robot appearance types on social interaction with a robot and service evaluation of a robot," *Archives of Design Research*, vol. 27, no. 2, pp. 81–93, 2014.

[15] J. Nakanishi, I. Kuramoto, J. Baba, K. Ogawa, Y. Yoshikawa, and H. Ishiguro, "Soliloquising social robot in a hotel room," in *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 2019, pp. 21–29.

[16] M. Mori, "Bukimi no tani [the uncanny valley]," *Energy*, vol. 7, pp. 33–35, 1970.

[17] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *CVPR*, vol. 1, no. 2, 2017, p. 7.

[18] J. Nakanishi, I. Kuramoto, J. Baba, O. Kohei, Y. Yoshikawa, and H. Ishiguro, "Can a humanoid robot engage in heartwarming interaction service at a hotel?" in *Proceedings of the 6th International Conference on Human-Agent Interaction*, 2018, pp. 45–53.

[19] M. A. Bayes, "Behavioral cues of interpersonal warmth." *Journal of Consulting and clinical Psychology*, vol. 39, no. 2, p. 333, 1972.

[20] Y. Chung-En, "Humanlike robot and human staff in service: Age and gender differences in perceiving smiling behaviors," in *2018 7th International Conference on Industrial Technology and Management (ICITM)*. IEEE, 2018, pp. 99–103.

[21] S. T. Fiske, A. J. Cuddy, and P. Glick, "Universal dimensions of social cognition: Warmth and competence," *Trends in cognitive sciences*, vol. 11, no. 2, pp. 77–83, 2007.

[22] H. S. Kattara, D. Weheba, and O. A. El-Said, "The impact of employee behaviour on customers' service quality perceptions and overall satisfaction," *Tourism and Hospitality Research*, vol. 8, no. 4, pp. 309–323, 2008.

[23] A. E. Lloyd and S. T. Luk, "Interaction behaviors leading to comfort in the service encounter," *Journal of Services marketing*, vol. 25, no. 3, pp. 176–189, 2011.

[24] K. Bergmann, F. Eyssel, and S. Kopp, "A second chance to make a first impression? how appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time," in *International Conference on Intelligent Virtual Agents*. Springer, 2012, pp. 126–138.

[25] F. E. Fernandes, G. Yang, H. M. Do, and W. Sheng, "Detection of privacy-sensitive situations for social robots in smart homes," in *2016 IEEE International Conference on Automation Science and Engineering (CASE)*. IEEE, 2016, pp. 727–732.