Design and Evaluation of Service Robot's Proactivity in Decision-Making Support Process

Zhenhui Peng
Hong Kong University of Science and Technology
Hong Kong, China
zpengab@connect.ust.hk

Yunhwan Kwon*
Dongguk University
Seoul, Republic of Korea
yunhwan1903@gmail.com

Jiaan Lu†
University of Michigan-Shanghai Jiao Tong Univeristy Joint Institute
Shanghai, China
lujiaan1768@sjtu.edu.cn

Ziming Wu
Hong Kong University of Science and Technology
Hong Kong, China
zwual@connect.ust.hk

Xiaojuan Ma
Hong Kong University of Science and Technology
Hong Kong, China
mxj@cse.ust.hk

ABSTRACT

As service robots are envisioned to provide decision-making support (DMS) in public places, it is becoming essential to design the robot's manner of offering assistance. For example, robot shop assistants that proactively or reactively give product recommendations may impact customers' shopping experience. In this paper, we propose an anticipation-autonomy policy framework that models three levels of proactivity (*high, medium* and *low*) of service robots in DMS contexts. We conduct a within-subject experiment with 36 participants to evaluate the effects of DMS robot's proactivity on user perceptions and interaction behaviors. Results show that a highly proactive robot is deemed inappropriate though people can get rich information from it. A robot with medium proactivity helps reduce the decision space while maintaining users' sense of engagement. The least proactive robot grants users more control but may not realize its full capability. We conclude the paper with design considerations for service robot's manner.

*This author was an intern in Hong Kong University of Science and Technology when the work was carried out.
†This author was an exchange student in Hong Kong University of Science and Technology when the work was carried out.

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300328

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; **User studies**; **Laboratory experiments**; *Interaction design theory, concepts and paradigms*;

KEYWORDS

Proactivity; Human-Robot Interaction; Robot Manner; Decision-Making Support

ACM Reference Format:
Zhenhui Peng, Yunhwan Kwon, Jiaan Lu, Ziming Wu, and Xiaojuan Ma. 2019. Design and Evaluation of Service Robot's Proactivity in Decision-Making Support Process. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3290605.3300328

1 INTRODUCTION

The use of service robots, which operate semi- or fully autonomously to perform services useful to humans [34], is becoming a trend in our daily life. For robots that work in the service industry and play the roles of shop assistants [25], receptionists [33], and museum guides [29], to name a few, providing information or recommendations to support human decision-making is a critical part of their job. Existing decision-making support (DMS) research mainly focuses on question understanding and content generation [1, 24, 38], but largely overlooks the design of robots' manner of service. For example, a robot assistant could wait for requests to come in and then react to the resulting situation. Alternatively, it could anticipate user needs and provide responses without being prompted. One can find similar types of actions in human assistants (e.g., waiter/waitress and shop assistants). These are characterized as different levels of proactivity in the occupational psychology literature [20, 36] and proven to affect socialization and performance at work [11]. It is thus interesting to explore if the proactivity of a robot would have a similar effect when offering DMS service.

Previous works suggest that humans are sensitive to robot's manner, i.e., way of behaving. For instance, people have better collaboration experiences with a robot that handles their disengagement in a submissive manner than the one acting dominantly [49]. In a collective decision task, robots expressing disagreement politely can defuse their human partners' frustrations and get people to change their minds [50]. Hence, it is possible that a decision-maker would feel and react differently when a robot provides its assistance proactively or reactively. We can design proper manner for robots in DMS services only if we have an in-depth understanding of these possible effects.

To study this possibility systematically, we need to have a formal definition and consistent implementation of robot proactivity in the course of DMS. Proactivity in occupational psychology is the anticipatory action that people initiate to impact themselves and/or others [20]. The closest concept to it in Human-Robot Interaction (HRI) research is autonomy – the ability to perform intended tasks based on current state and sensing without human intervention [15]. It emphasizes primarily on initiating actions automatically based on what is happening in the collaboration [18, 40]. While previous works show that robot autonomy can cause perceptual differences in users [2, 49], it mainly concerns the current state of the machine. Anticipation, another important element of proactivity that emphasizes assumptions of what is going to happen, is not incorporated into the autonomy construct.

In this paper, we formulate three levels of robot proactivity (*high, medium, low*) in two dimensions – anticipation and autonomy – based on the occupational psychology and autonomy literature. We then design associated behavior policies in the context of a structured decision-making support process. To evaluate the impact of robot proactivity on decision-makers perceptions and behaviors, we conduct a within-subject, Wizard-of-Oz experiment with 36 participants. The participants reflect that the high-proactivity robot is the least appropriate though it can provide rich information, while the medium-proactivity robot is more helpful as it helps to narrow down choices. Thematic analysis on user behaviors further reveals that the participants can adjust their turn-taking behaviors to the robots' manner, and they have more control over the conversation with the low-proactivity robot by actively making requests. However, they engage better with the medium-proactivity robot, with sufficient opportunities to express their thoughts and feelings. Our work provides insights into designing a proper way of behaving for robots serving in DMS context, and adds to the understanding of how people perceive and interact with service robots of different manners.

2 RELATED WORK

Defining Robot's Proactivity

Derived from the definition of proactivity in occupational psychology [20, 36], service robot's proactivity can be defined as the anticipatory action that robots initiate to impact themselves and/or others. The definition indicates three elements of robot's proactivity: 1) anticipation; 2) initiation of action; 3) target of impact. The first element, anticipation, is the robot's assumption of what a human is going to do. There are many possible ways, e.g., gaze [23], body orientation [22], trajectory [42], etc., which can build up robot's anticipation of a human. For example, Koppula et al. used a rich graphical model based on object locations and human poses to anticipate what a human partner would do next, so that the robot could perform an anticipatory action, e.g., open the door for the human [30]. Bohus et al. made use of vision processing (e.g., face detection, distance) and speech recognition to determine if the visitors are going to be engaged or not in a direction-giving context [6].

The second element, the initiation of action, is more about system autonomy in the robot's context, as suggested by many other works about designing robot's proactive behaviors [2, 9, 55]. According to the amount of human input and the authority of robot, the levels of autonomy in HRI can be divided into 10 levels [3, 40, 44]: 1) robot offers no assistance; 2) robot offers a complete set of action alternatives; 3) robot narrows the selection down to a few choices; 4) robot suggests a single action; 5) robot executes that action if human approves; 6) robot allows the human a limited time to veto before automatic execution; 7) robot executes automatically then necessarily informs the human; 8) robot informs human after automatic execution only if human asks; 9) robot informs human after automatic execution only if it decides to; 10) robot decides everything and acts autonomously, ignoring the human. A proactive robot is usually designed to have a higher level of autonomy to take the initiative [35]. For example, in Fink's work [13], a proactive robot was designed to initiate the interaction to motivate the children to tidy up their toys, while the reactive robot waited until the children take actions first. Also in Baraglia's work [2], the most proactive robot took the initiatives to help human whenever it could help, while the least proactive robot only helped human when being requested.

The last element, target of impact, is the object of robot's anticipatory actions. In human-robot interaction (HRI), the intended target would be the human partner [19].

This paper focuses on a particular type of HRI scenario – supporting human decision-making, where robot's goal is to fulfill its capability of providing information interactively to improve the quality of user's decisions [27]. In such a setting, a highly proactive robot builds strong anticipation of human user's information needs and takes the liberty of making recommendations. In contrast, a less proactive robot is more conservative in making assumptions and intervening. In any case, DMS robots should behave in a user-friendly and socially acceptable manner [12].

Manner of Decision-Making Support Systems

Previous studies have explored different ways for a decision-making support (DMS) system to perform its job, i.e., providing information or recommendations to people. For example, to improve the current recommendation, Li et al. suggested that the recommendation system can use system- or user-initiated critiquing methods to seek feedback from the user [10]. The system-initiated critiquing system automatically generates knowledge-based critiques (e.g., "Different Manufacture, Lower Processor Speed and Cheaper Cameras"), which could familiarize the user with the product domain but may fail to match user's specific criteria. In contrast, the user-initiated critiquing system lets user make self-motivated critiques from a complete set of options, which allows for a higher level of user control with the drawback of more user's efforts. Similarly in a fantasy baseball game, Solomon found that giving more users control over the DMS system (i.e., let them customize it) could make the recommendations more acceptable regardless of their accuracy [47].

Woiceshyn et al. tested the system-initiated approach on a robotic DMS system, which can actively ask for user preference, give and update recommendations according to user's explicit feedback (accept or reject) [52]. Their experiment showed that the system was easy to use and effective. Rau et al. found that a robot with higher autonomy, which gave recommendations before the user carried out any action, would have greater influence on people's decision-making in a sea survival decision task, compared to the robot that gave opinions only after the user made a decision [40]. In a more real-world setting, Shiomi et al. found that the robot that directly recommended one choice (autonomy 4) could increase the number of specific store coupons that people finally chose, compared to the robot that provided all available candidates (autonomy 2) [46]. Note that all of the aforementioned studies primarily focused on the autonomy aspect of a DMS robot's way of acting. In this work we also incorporate anticipation into the DMS robot's manner design. Different from previous works that emphasized mainly on decision outcomes, we are interested in how robot's proactivity level (anticipation + autonomy) would affect user experience in the decision-making process, a critical determinant of the success of DMS systems [39].

Effect of Robot's Proactive Manner

To achieve more natural and efficient human-robot interaction (HRI), researchers in HRI have spent many years on designing and evaluating the proactive manners of robot. For example, to serve people, the robot needs to properly approach and initiate the interaction with people [16, 26, 42, 45]. Learned from observation of assistant's behavior, Kato et al. [26] proposed a robot strategy that can exhibit "availability" through body orientation and gaze before approaching people. Their field studies in a shopping mall showed that their strategy is less intrusive compared to the "passive waiting" and the "simply-proactive" strategy. In a handover task, Huang et al. derived different robot coordination methods from human-human collaboration, and found that there was a tradeoff between team performance and user experience using different methods [22]. Compared to reactive (wait for user's completion first) and adaptive (wait and slow down to adapt to user's availability) coordination methods, the proactive coordination (always present next object) can significantly improve team performance but impair users' perceptions of the robot. The tradeoff also appears in a simulated Urban Search and Rescue task, in which the researchers found that people preferred the robot that can provide proactive support. But human cognitive load was also increased with the proactive robot [55]. Although these studies were not conducted in DMS context, it is reasonable to hypothesize that robot's proactivity would have an effect on decision makers' perception and interaction behaviors.

3 DESIGN OF DMS ROBOT'S PROACTIVITY

Principles for Designing Different Proactivity

Based on the robot's anticipation and levels of autonomy [20, 40], we derive the principles for designing three levels of robot's proactivity:

High-proactivity: the robot makes strong assumptions (e.g., users need help, like popular items, etc.) and actively offers it; the robot automatically takes action with very small amount of user input during the interaction (autonomy 6-8).

Medium-proactivity: the robot makes some assumptions and lets the user verify them; the robot needs some amount of user input to take automatic actions during the interaction (autonomy 3-5), e.g., providing limited choices like "A or B".

Low-proactivity: the robot makes no assumption and needs the users to explicitly tell what they want; the robot only responds upon request and needs a large amount of user input (autonomy 1-2), e.g., providing a complete set of choices.

Associated Behavior Polices in DMS Settings

To show how these design principles for robot's different proactivity can be applied to decision-making support (DMS) settings (e.g., restaurants, airports, banks, stores, etc.), we structure a DMS process (adapted from [10, 52]) and design associated behavior policies for the robots as shown in Figure

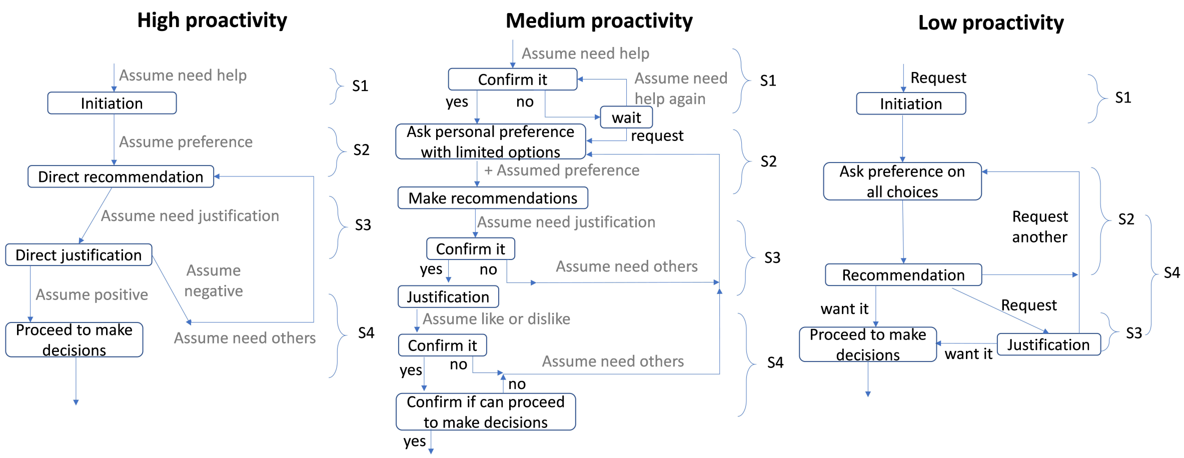

Figure 1: Robot's behavior policies with different proactivity levels in a structured decision-making support process.

1. In the rest of this section, we use shopping as an example to showcase the robot's behaviors under different policies.

S1. Initiation: To initiate the interaction, either the robot or the user should take actions to start the conversation. **1)** The *high-proactivity* robot assumes that the user needs help and actively offers it, e.g., "Hi, have troubling finding things you like? Let me show you some of our popular items!". **2)** The *medium-proactivity* robot also assumes that the user needs help but confirms it first, e.g., "Have trouble deciding? Would you like to know about our best collection?". Only with user's approval will the robot move to the next step, otherwise it waits for the next trigger event. **3)** The *low-proactivity* robot makes no assumption and only offers help, e.g., saying "How can I help you?" upon request.

S2. Preference elicitation and recommendation: After initiation, the robot gives recommendations to the customer. **1)** The *high-proactivity* robot assumes knowing user's preferences (e.g., popularity) and does not ask for such information; instead, it directly makes recommendations, e.g., "Check out this popular item here!". **2)** The *medium-proactivity* robot assumes knowing some preferences (e.g., popularity) but asks for personal preferences by providing limited options first, e.g., "Which [attribute] do you want, [option 1] or [option 2]?". It then makes recommendations based on assumed preference and personal preference. **3)** The *low-proactivity* robot has no assumption of user's preferences and asks for it by providing all choices, e.g., "We have [all choices], which one do you like?". It then makes recommendations based on user's responses.

S3. Justification: After presenting the recommended items, the robot might need to justify its recommendations. **1)** The *high-proactivity* robot assumes that the customer wants justification and provides it directly, e.g., "The special point of this item is [...]". **2)** The *medium-proactivity* robot assumes that the customer wants justification but confirms it first, e.g., "If you are interested in this item, I can tell you more about it. Otherwise, I can give you another recommendation". If the customer is not interested, it asks for user's preferences again to provide another recommendation. **3)** The *low-proactivity* robot makes no assumption about what the user wants to know, and only justifies the item upon request.

S4. Feedback seeking: After justification, the robot needs to seek feedback from the user to either proceed or revise recommendations. **1)** The *high-proactivity* robot tells from the user's reaction (e.g., silent duration, facial expression, gaze, etc.) to assume if the user likes the recommended item or not. If assuming that the customer likes it, the robot proposes to proceed further, e.g., "Seems that you are quite interested. I have the order form ready for you to fill out". If assuming that the user dislikes it, the robot proactively provides another recommendation, e.g., "It seems like you are not very happy with this item, let me show you something else". **2)** The *medium-proactivity* robot assumes that the user may like it or dislike it, and confirms it, e.g., "Do you like it?". If the user likes the recommended item, it confirms whether they can go further, e.g., "Ok, if you want to buy this pair, please fill out the order form". Otherwise, it asks user's preferences again to provide another recommendation. **3)** The *low-proactivity* robot makes no assumption about the user's feedback, and waits for the user to explicitly tell it. If user dislikes the item and asks for another one, it asks for user's preferences again to provide another recommendation. If the user explicitly says that he/she wants the item, it proceeds to make decisions.

4 EXPERIMENT

To verify our service robot's proactivity design in the DMS context and to evaluate its impact on user perception of the

Table 1: Persona and shoes for each task.

Persona	Shoe type	Color	Occasion
Men	Oxfords	Black or Brown	Dress or Casual
Women	Heels	Black or Beige	Dress or Casual
Teens	Sneakers	Black or White	Skate or Running

robot and interaction behaviors, we conduct a within-subject, Wizard-of-Oz experiment with 36 participants.

We simulate an offline shoe shopping scenario and customize the robot behaviors (described in the section Design of Robot's Proactivity) by replacing "item" by "shoe". In the experiment, each participant interacts with three versions (*high-*, *medium-* and *low-proactivity*) of robots separately to pick a pair of shoes that is suitable for a specific persona as a gift. We counterbalance the task assignment and the order of three robot conditions to minimize the potential order effect.

We create information (e.g., portrait, gender, age, work, short bio, goals, motivations, etc.) of the persona and corresponding shoe collections for each task, i.e., *Men's Oxfords, Women's Heels* and *Teens' Unisex Sneakers* (note: the participant can either pick a girl or a boy persona in the *Teens* task as the shoes are unisex). We collect relevant shoes data from Zappos shoe website [1]. For each task, we prepare 32 pairs of popular (sorted by "Best Sellers") shoes which can be classified based on two attributes: color and occasion (two options for each attribute, as shown in Table 1). To minimize potential bias introduced by price and brand, we keep the price of the shoes within a reasonable range (e.g., $79.95 - $99.95 for Men's Oxfords), and blur the brand signs. To get participants serious about their tasks, we ask the them to justify their final choices, showing how they resonate with real experiences. Note that similar to [53, 54], there is no right or wrong decision, but only participants providing valid reasons can enter a lucky draw for two extra coupons.

Robot and Wizard-of-Oz Study Design

In this study, we use the Pepper from Softbank Robotics [2], a robot widely adopted in the service industry, as a shop assistant to support people's shoe purchasing decisions. This humanoid robot is 120 cm tall and has a 246×175 mm tablet in the front. We use the built-in packages of the robot to generate its speech in a gender-neutral voice. It can display natural body movements and present a picture of a recommended item on its tablet during the DMS process. We build a shoe-related knowledge base for the robot, which consists of the popularity (i.e., # of "like"), properties (e.g., insole, toe style, material), special feature (Table 2) of each pair of shoes in our collection. The robot may make recommendations based

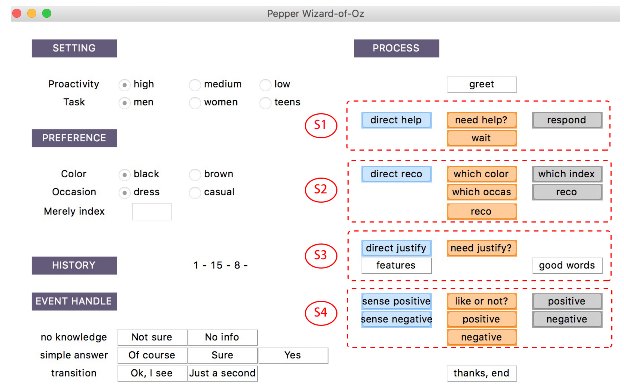

Figure 2: The wizard interface. S1-S4 shows the robot's behaviors in the four steps of the DMS process (blue: *high*, orange: *medium*, grey: *low*, white: common). Setting different proactivity and task will hide irrelevant buttons and customize the contents of the relevant buttons.

on the name, popularity, specification of two attributes (color and occasion), special feature of the intended shoes. We prepare multiple scripts (modified from shoppers' reviews) for each special feature, as well as some general good words to advertise. We carefully test all the scripts in a pilot study to make sure that they can be clearly heard and understood from the robot (script samples in Table 2).

We adopt a Wizard-of-Oz approach for the evaluation to avoid possible interference of technological pitfalls, since state-of-the-art technologies still fail to perfectly understand human speech and intent in real-world settings [4, 43]. The first author acts as a wizard to anticipate user intention and control subsequent robot actions strictly based on the proposed behavior policies in each task (Figure 1). Similar to Shamekhi's method [43], we use a constrained Wizard-of-Oz protocol [41] with a small set of intents which are elaborated on Figure 1 (e.g., need help, need justification).

The human wizard tries to infer participants' intentions from their speech, head pose, eye gaze and facial expressions, and then trigger associated robot responses through an easy-to-use interface (Figure 2). For example, in the initiation step, if noticing that the participant turns to the robot, has a hesitant or confused look on the face, or keeps browsing for some time without inspecting any particular item, the wizard would assume that the participant needs help. Consequently, the wizard would instruct the robot to either offer help directly (high proactivity), ask for permission to intervene (medium), or wait for the participant's explicit help seeking signal (low). Pilot studies reveal that participants may skip some steps in the structured DMS process, e.g., directly asking for a specific pair of shoes after the first-round justification in high- and medium-proactivity tasks. In

[1]https://www.zappos.com/c/shoes
[2]https://www.softbankrobotics.com/emea/en/robots/pepper

Table 2: Script samples of special features and general good words that the robot uses for justification.

Feature & good words	Scripts Sample
Suitable for walking	"They are very suitable for walking. If you have to wear shoes for a long time, these are absolutely a good option. I think they will be one of the most comfortable shoes you have ever owned."
Stylish and fashionable	"They are very stylish and fashionable. They're also beautiful, with a very nice shape and popular toe style. I think they will keep you a fashionable person in the public."
Water proof, easy to clean	"They are water proof. Even in harsh rainy conditions, you can wear them walking around streets for eight hours a day. I think they are very suitable for those who need to walk in dirty ground."
Latest design	"These are the latest shoes available. They are light weight, incredibly comfortable, and trendy. I think you can never go wrong with this new design."
Best materials	"The material of these shoes is very special. The soft insole makes the shoes comfortable for long wear. I think they are also a very classical shoe, which can last than 20 years."
General good words	"These will be your favorite shoes for sure." "Buy this pair, you will never regret it."

these cases, the robot will first satisfy their requests and then resume its policies in the following turns. For instance, the high-proactivity will still assume the user needs justification and directly justify the requested shoe. We also prepare additional scripts, i.e., no knowledge (e.g., "sorry, it is not written on the label"), simple answer (e.g., "of course") and transition words (e.g., "ok, I see"), for handling some unexpected events to ensure the smoothness of the whole HRI experience.

Hypotheses

Previous works suggest that the highly proactive robot behaviors can impair users' perception of the robot, making the robot less appropriate and less likeable [22, 49]. And in Human-Human interaction some moderate proactive behaviors could trigger higher rate of socialization and higher job performance [11]. Therefore, we hypothesize that:

H1. Compared to the *medium-* and *low-proactivity* robots, the *high-proactivity* robots are perceived to be *(H1a)* significantly less appropriate, *(H1b)* less polite, *(H1c)* more controlling and *(H1d)* more interrupting.

H2. Compared to the *high-* and *low-proactivity* robots, the *medium-proactivity* robots are perceived to be *(H2a)* significantly more helpful. Participants *(H2b)* depend significantly more on their recommendations, *(H2c)* significantly get more information from the medium-proactivity robots, and *(H2d)* prefer to be served significantly more by the medium-proactivity robots in the future.

We measure these aspects (adapted from [31, 39, 48, 49, 51]) and robot's proactivity on a standard 7-point Likert scale (1 - strongly disagree, 7 - strongly agree).

Participants

Thirty-six students (P1-P36, 18 females and 18 males, balanced across the counterbalanced conditions) from the local university participate in our experiments. Participants study a diverse range of fields, and their ages range 18-30 ($M = 23.75, SD = 2.52$). Their average familiarity with robots

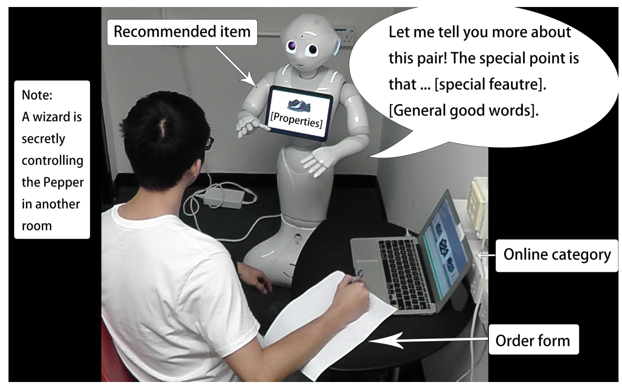

Figure 3: The *high-proactivity* robot is justifying its recommendation without the participant's request.

is 3.14 (SD = 1.43), with 1 for *no experience at all* and 7 for *a lot of experience*. Thirty-two of them report that they have experience of interacting with physical or virtual conversational robots. All participants have a TOEFL score higher than 88 or an IELTS score above 6.5, meaning that they have no problem communicating with the robot in basic English.

Procedure

After obtaining consent from the participants, we introduce the procedure of the experiment. The participant is informed that the physical store is rather small and does not have everything in the store locally, and that he/she can browse the online category first. In each of the three sessions, the participant first reads information of a specific persona in an A4 paper. Then the participant enters the study room and the robot invites the participant to sit down. On the table, a laptop (Mac Air 11.6", Intel CPU 1.6GHz) provides high-quality images of each pair of shoes and the wizard in another room uses it to monitor the process (Figure 3). The participant starts to browse the shoes on the computer with the support of the robot in the decision-making. Once the participant makes a final decision and fills out the order form on the table, he/she can call "Pepper" and the robot will respond "Thank you! It has been a pleasure to serve you. Have a nice day!" to end the interaction as a human assistant would do. At the end of each session, we ask the participant to write down the reason why his/her final choice is suitable and fill out a questionnaire to rate their perception of the robots on a 7-point Likert scale. Upon the completion of the three sessions, we conduct an in-depth interview with the participant to find out more about his/her feelings regarding the robot's proactivity. After debriefing, each participant receives some compensation. The whole process lasts for around 40 minutes for each participant.

Behavioral Analysis Method

To evaluate how the robot's proactivity influences user behaviors during the interaction, we record the video of each session with user's consent. We conduct a thematic analysis [7] on the participants' behaviors during their interaction with the robots. Two researchers first familiarize themselves by watching 9 videos of 3 participants to identify interesting events occurred in the human-robot interaction. Through discussion, we narrow it down to several points related to the users' turn-taking behaviors in the interactions, content of their turns, and their explicit attitudes shown to the recommended items. Then the two researchers openly code 30 videos of 10 participants. During this process, the two researchers meet regularly to compare, discuss and refine the codes, grouping codes into potential themes. After several rounds of discussion and theme refinement, we generate an embryonic form of code book and apply it to all the videos. Next, we carefully review the themes, including combining some codes into several categories, splitting the codes that can be put into different themes and discarding irrelevant codes. Finally we name the themes and categories, as well as count their occurrences in different conditions.

5 RESULTS

Since the shoes in the tasks (Table 1) are gender-specific, we want to inspect the possible effect of participants' gender on their decision-making experience. We first run a two-way mixed ANOVA (practivity level as within-subject and gender as between-subject) on each of the quantitative measures of the participants' perceptions. Neither the main effect of gender nor the interaction effect (proactivity × gender) is significant. We conduct another two-way mixed ANOVA to check for possible order effect. Similarly, the result suggests that neither the main effect of the order nor the interaction effect (proactivity × order) on the quantitative measures is significant. Thus in the following statistical analysis we treat the robot's proactivity level as the only independent variable, and conduct a one-way repeated measures ANOVA on the statistical results. For each ANOVA, the assumption of equal variance is confirmed by the Macuchly's test of sphericity or otherwise adjusted using Greenhouse-Geisser [17]. In the rest of this section, we summarize the statistical analysis, behavioral analysis and interview results, in terms of perceived robot's appropriateness, helpfulness and user behaviors during the interaction.

Manipulation Check

The manipulation check for robot's proactivity shows that the manipulation is effective ($F(2, 70) = 59.49, p < .01, \eta^2 = 0.74$). Bonferroni post-hoc test confirms that all pairwise comparisons are significantly different ($p < .01$). The high-proactivity robot is perceived to be the most proactive ($M = 5.64, SD = 1.10$), followed by the medium-proactivity robot ($M = 4.58, SD = 1.46$) and then the low-proactivity robot ($M = 3.03, SD = 1.50$).

Perceived Appropriateness

The left-hand side of Figure 4 shows the statistical results of user perception of the robots in terms of appropriateness. In general, participants feel that the high-proactivity ($M = 4.72, SD = 1.28$) robot's behaviors are significantly less appropriate than both the medium- ($M = 5.58, SD = 1.13; p < .01$) and low-proactivity ($M = 5.67, SD = 1.15; p < .01$) robots; $F(2, 70) = 8.51, p < .01, \eta^2 = .20$; Bonferroni post-hoc test; ***H1a*** accepted. Post-study interviews suggest that the high-proactivity robots are often perceived to be in pursuit of the store's profit and they somewhat intrude into the participants' decision space, which is not a user-friendly and socially appropriate manner. *"It feels like it (high) is selling something"* (P29, male, age: 24). *"I can't make satisfied choice as it (high) doesn't give me enough time"* (P4, female, age: 24).

Participants also feel that the politeness of different roots is significantly different ($F(2, 70) = 4.19, p < .05, \eta^2 = .11$). Bonferroni post-hoc test shows that the high-proactivity robot ($M = 5.31, SD = 1.33$) is significantly less polite than the medium-proactivity one ($M = 5.97, SD = .94; p < .05$). However, the difference between high- and low-proacitivity ($M = 5.83, SD = 1.32$) robots is not significant. ***H1b*** is partially accepted. The participants really appreciate the politeness of our robots, especially the medium-proactivity one. *"The second (medium) one is more polite as it doesn't use so many words as the first (high) one"* (P1, male, age: 26).

Furthermore, participants suggest that the high-proactivity robot ($M = 4.92, SD = 1.44$) takes significantly more control of the conversation than both the medium- ($M = 3.97, SD = 1.52; p < .01$) and low-proactivity ($M = 2.86, SD = 1.57; p < .01$) ones; $F(2, 70) = 20.66, p < .01, \eta^2 = .37$; ***H1c*** accepted. Bonferroni post-hoc test shows that the medium-proactivity robot is also significantly more controlling than the low one ($p < .01$). Similar effects are found in the participants' perception of being interrupted. The participants feel that they

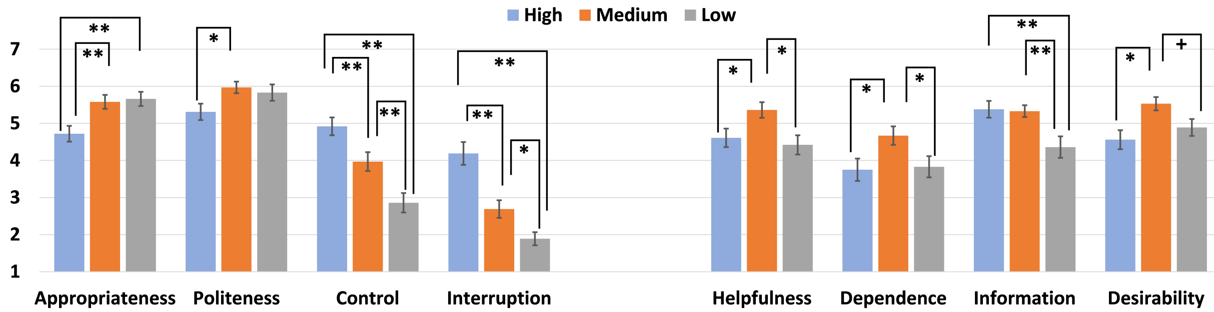

Figure 4: Means and standard errors of the user perception of the robots in terms of appropriateness (left) and helpfulness (right) on a 7-point Likert scale ($+: .05 < p < .1, *: p < .05, **: p < .01$).

are interrupted by the high-proactivity ($M = 4.19, SD = 1.85$) robot significantly more often than by the medium- ($M = 2.69, SD = 1.43; p < .01$) and low-proactivity ($M = 1.89, SD = 1.06; p < .01$) ones; $F(2, 70) = 24.40, p < .01, \eta^2 = .41$; ***H1d*** accepted. The sense of being interrupted very often could be the main reason for the inappropriateness of the high-proactivity robot. *"I don't like to be disrupted by them (high, medium). The low one is better"* (P11, male, age: 25).

Perceived Helpfulness

Regarding perceived helpfulness (right-hand side of Figure 4), participants generally think that the medium-proactivity ($M = 5.36, SD = 1.27$) robot is significantly more helpful than both the high- ($M = 4.61, SD = 1.50; p < .05$) and low-proactivity ($M = 4.42, SD = 1.56; p < .05$) ones; $F(2, 70) = 5.43, p < .01, \eta^2 = .13$; Bonferroni post-hoc test; ***H2a*** accepted. The participants also feel that they depend significantly more on the recommendations from the medium-proactivity ($M = 4.67, SD = 1.49$) robot than from the high- ($M = 3.75, SD = 1.81; p < .05$) and low-proactivity ($M = 3.83, SD = 1.71; p < .05$) ones; $F(2, 70) = 3.53, p < .05, \eta^2 = .09$; Bonferroni post-hoc test; ***H2b*** accepted. Many participants comment that the medium-proactivity robot can give more proper recommendations and help reduce decision space interactively. *"(Medium) It is better to learn my preference first. Otherwise, the recommendations are useless"* (P22, female, age: 22). *"It is more convenient that it (medium) can automatically help me narrow down choices. It is more efficient"* (P25, male, age: 22).

However, in some cases, the high- or low-proactivity robots are of better assistance to the users. For example, the high-proactivity robots are appreciated when the participants are not familiar with the items, while the low-proactivity robots are more efficient when the participants already have something in mind. *"If I am not familiar with this type of item, I would like the robot to proactively introduce it first"* (P26, female, age: 23). *"I have almost made up my mind and want to confirm it. It shouldn't be so proactive"* (P17, male, age: 24).

The differences of the rating *"I get a lot of information from the robot"* among three cases are also significant ($F(2, 70) = 9.43, p < .01, \eta^2 = .21$). The participants perceive that they get significantly more information from the high- ($M = 5.39, SD = 1.36; p < .01$) and medium-proactivity ($M = 5.33, SD = .96; p < .01$) robots than from the low-proactivity ($M = 4.36, SD = 1.74$) robots; Bonferroni post-hoc test. However, the difference between the high- and medium-proactivity robots is not significant. ***H2c*** is partially accepted. Although many participants find the manner of the high-proactivity robot not as apt as that of the other robots, they do agree that they can get rich information from it. Some of them actually appreciate that the robots can provide information without being explicitly requested. *"It's great that the robot proactively provides more information when I am hesitant. It broadens my mind as there are some points I didn't consider"* (P32, female, age: 23). In contrast, many participants report that they can not easily get much out of the low-proactivity robot because they do not know its capability. *"I don't know what it (low) can do. It should let me know it has such functions"* (P34, female, age: 20).

Overall, participants prefer to be served by the medium-proacitivty robot in the future, significantly more than the other two ($F(2, 70) = 7.29, p < .01, \eta^2 = .17$). The Bonferroni post-hoc test suggests that people prefer significantly more the medium- ($M = 5.53, SD = 1.08$) than the high-proactivity ($M = 4.56, SD = 1.54; p < .05$) robots. However, the difference between medium- and low-proactivity ($M = 4.89, SD = 1.37; .05 < p < .1$) robots is only marginally significant. ***H2d*** is partially accepted. The participants who prefer the low-proactivity robot argue that it can grant them more control over the conversation and flexibility of seeking information. *"I want to control the conversation. It is enough that it has the functions to provide me information"* (P21, male, age: 21). *"I can check the information by myself. The robot doesn't need to explain it"* (P30, female, age: 22).

Table 3: Summary of users' behaviors during interaction with different robots (average occurrences, SD is in the parenthese). Note that an event may contain multiple codes and thus falls into different themes.

Theme	Category	Code example	High	Medium	Low
Turn-taking behaviors	1. Initiate the turn	(Robot is waiting) "I have a friend [...] do you have any recommendations?" (After robot justification) "Do you think it is suitable for a single man?"	2.3(2.02)	3.0(2.62)	6.2(5.09)
	2. Competing for the turn	(Robot proposes to justify or is justifying the shoes) "How about No. 32", "No, I don't want this one", "Let me check"	2.5(2.09)	2.0(1.84)	0.7(0.97)
Purpose of users' turns	3. Make requests	"Could you recommend me another pair?" "Show me No. 12" "Can you give more information?" "Please tell me more about it"	2.5(2.73)	3.7(3.09)	4.5(3.65)
	4. Ask questions	"Do you think it is suitable for a very busy woman?" "Which pair of shoes is most fashionable?"	0.8(1.42)	1.0(2.16)	1.5(2.22)
Attitudes to recommended item	5. Positive	(Robot gives recommendations) "Okay I like this pair" "Yeah, thank you!" (Robot proposes to justify or is justifying the shoes) "Okay" "Yes, show me"	2.2(2.25)	3.3(2.35)	0.8(1.02)
	6. Negative	(Robot gives recommendations) "No, next!" "I don't like this one" (During justification) "Probably no" "(ignore robot) give me another pair"	1.8(1.65)	3.1(3.08)	0.8(0.95)

User Behaviors

The difference among the times that the participants interact with the high- ($M = 273.9.5s, SD = 113.6s$), medium- ($M = 254.5s, SD = 114.6s$), and low-proactivity ($M = 254.0s, SD = 119.0s$) robots is not significant; repeated measures ANOVA, $F(2,70) = .55, p = .58, \eta^2 = .015$. We summarize the final themes through behavioral analysis in terms of *Turn-taking behaviors*, *Purpose of users' turns* and *Attitudes to recommended item* (Table 3). We also count the occurences of codes in each category ($1-6$ as shown in Table 3) of the theme with different robots. An event may contain multiple codes in terms of different aspects, and thus falls into different themes. For example, the event *(The robot is justifying the shoe) "No, please show me shoe No.32"* (P27, male, age: 30) falls into categories 2, 3 and 6 with different themes.

Turn-taking behaviors. In general, people interacting with the robots with low proactivity initiate the turn significantly more often ($F(1.58, 55.46) = 18.79, p < .01, \eta^2 = .35$) and compete for the turn significantly less often ($F(1.64, 57.28) = 13.70, p < .01, \eta^2 = .28$) than robots with high or medium proactivity. These results show that the participants can adapt to the robot's capabilities and adjust their turn-taking behaviors accordingly [37]. These behaviors are also consistent with the perceived control and perceived interruption of different robots (Figure 4). Participants actually lead the conversation with low-proactivity robots by initiating the turn more often. They freely tell the robots what they need and actively ask some questions, e.g., *"(Robot just gave recommendation) do you have any other color? I think they are a little bit simple"* (P11, male, age: 25), *"(Robot just gave justification) Are they comfortable for playing football?"* (P12, female, age: 22). While in the high-proactivity cases, the participants compete for turns more often, especially when they disagree with the robot, e.g., *"(Robot is justifying) no, I don't think so, he stays at home ..."* (P11, male, age: 25). Compared to human assistants, some participants feel less pressure while competing for the turn with the robot. *"I don't need to care about its feelings. I can interrupt or ignore it if I want"* (P5, male, age: 24).

Purpose of users' turns. The occurrences of making requests in the users' turns are significantly different among three cases ($F(2,70) = 5.59, p < .01, \eta^2 = .14$). Bonferroni post-hoc test reveals that the participants request significantly more often from the low- than from the high-proactivity robots ($p < .01$). The participants interacting with the low-proacitivty robots usually give concrete descriptions of what they want, e.g., *"another point is that my friend likes hanging out with his friends, and also likes recording interesting moments..."* (P11, male, age: 25), while the participants interacting with the high- and medium-proactivity robots normally just give simple commands to make requests. An interesting thing is that in all conditions the participants tend to use words like "could", "please" to formulate the requests, which matches the negative politeness that aims to keep their social distance with the others [8]. There are no significant differences in the occurrences of asking questions in the users' turns among three conditions ($F(2,70) = 1.84, p = .17, \eta^2 = .05$). Nevertheless, it is interesting that some participants ask specific questions in their turn, actively seeking the robot's opinions, e.g., *"are they suitable for walking for a long time?"* (P18, female, age: 23). Their behaviors suggest that the participants were treating the service robot like real human assistant rather than a tool [14, 32]. It implies the need to design social intelligence in service robots, following the norms in interpersonal interactions [12]. *"I treat it more as a human and I want interaction. It uses the same attitude to serve human and gives answers quickly"* (P29, male, age: 24).

Like human shop assistants, robots may receive requests or questions outside the scope of their knowledge, e.g., "which age is suitable for those shoes?" (P22, female, age: 22). Such incidents occur more often in the low-proacitivity cases (total occurrence: 26) than in the medium (14) and high (10) conditions. It is possible that without the robot revealing possible dimensions of its decision space, people may have trouble asking the "right" type of questions that the robot can handle. As a result, *"...it looks like it (low) can't provide satisfying service"* (P15, male, age: 22).

Attitudes to Recommended Item. In total (positive and negative), the occurrences of users' explicit attitudes to the recommended items during the HRI are significantly different ($F(1.48, 51.92) = 19.67, p < .01, \eta^2 = .36$). Bonferroni post-hoc test further shows that the participants explicitly express their thoughts about the recommended items significantly more to the medium-proactivity robots than to the high- ($p < .05$) and low-proactivity ($p < .01$) robots. People interacting with the low-proactivity robots usually do not explicitly tell their feelings, but just initiate the next turn by making other requests or asking questions. Compared to the high-proactivity robots, participants engage with the medium-proactivity robots more actively. They actively express their positive or negative attitudes to the recommended items, updating robot's anticipation and steering robot's next move. In fact, some participants really enjoy engaging with the medium-proactivity robots, probably because in this case they can have some guidance during the interaction. *"It (medium) asks me detailed preferences, which can help me be clear in my mind. And I can modify my preferences by expressing my current feelings"* (P30, female, age: 22).

In summary, our design of robot behavior policies can successfully convey different levels of robot's proactivity to the participants. Decision-makers do perceive and react differently to robots operating under various levels of practivity. Although the high- and low-proactivity robots can be useful in some cases, the decision-makers generally prefer the medium-proactivity robot that can interactively and proactively provide them needed information with their permissions. These results indicate that the design of the robot's manner is critical for providing satisfying services.

6 DISCUSSION

Design Considerations for Service Robot's Manner

We derive several design considerations for service robot's manner from our experimental findings, which signifies the trend of developing human-aware intelligent systems [9].

Robot Should Maintain a Mental Model of Human. To achieve effective collaboration, service robots should model user's mental state such as goal, preference and knowledge [9]. In our experiment, the low-proactivity DMS robot suffers from the lack of a mental model of its user. *"It (low) is too passive, giving me the feeling that it doesn't devote itself to the work and knows nothing"* (P21, male, age: 21). Decision makers expect a DMS robot to infer their information needs based on some reasonable assumptions. However, instead of directly offering information based on the anticipated needs, in most cases it is deemed more considerate for the robot to verify its mental model prior to taking any action. This can mitigate the possible negative consequences of acting upon an incomplete mental model with uncertain information. Interestingly, once users approve the robot's mental model, they seem to trust the robot's judgment more and be more willing to let the robot take the initiative in the following process. *"The robot solved my first question, and it proactively provided information to other related issues that I might be interested in. Really like it"* (P34, female, age: 20).

Robot Should Express Its Capability. Some participants complain that they do not know what the robot can do during the interactions, especially in the low-proactivity condition. Lacking knowledge of the robot's capacity may lead to incorrect expectations. Sometimes users may think too highly of the robot and ask questions beyond its scope of knowledge. In some other cases, users underestimate the robot and fail to make full use of its service. Both situations can cause frustration and dissatisfaction. We suggest that robot should interactively help human users obtain an accurate mental model of its capability. For example, if the user fails to ask a "correct" question for the first time, the robot could 1) show uncertainty through some cues (e.g., "um" [5]), head motion [49]); 2) explain the cause of its incapability [28]; and/or 3) prime users about questions that it can handle [31].

Robot Behavior Policy Should Be Adaptive. Some participants suggest that their preferences for different levels of robot proactivity are context dependent, e.g., familiar with the items or not, in a hurry or not, knowing the robot well or not, etc. It is thus necessary for robots to adapt their behavior policies to the changes of context. For example, in the shopping scenario if a user keeps asking for some basic information of the items, the robot may infer that he/she is not very familiar with the options and could offer to explain more about possible candidates. Besides, robots should be sensitive to users' emotional reaction to its behavior. For example, when a service robot approaches a user and proposes to help, if the user says "no" or expresses impatience or avoidance, it may be better for the robot to reduce its level of proactivity. In such a case, waiting for users' request is more likely to result in a satisfying user experience.

Limitation

Our work has several limitations. First, we only use shoe shopping scenario to showcase the common decision-making support process in our study, while more experiments in different scenarios (e.g., direction-giving, food ordering, etc.) are needed to evaluate the design of robot's proactivity. Second, we conduct our study in a controlled environment. In a real-world setting, the task can be less structured or confined, with a bigger decision space and possible interference from the background. For example, different aspects of interaction dynamics like timing [49] and tone [21] might influence the interaction, while in our study we set the delay between corresponding actions and anticipation to zero and kept a consistent polite tone. Third, we are aware of the possible moderating effects of various user characteristics but only examine gender in this paper. For example, people from various cultural backgrounds may assess the appropriateness of robot manner using different social standards and norms. The elders and children may have different expectations of a service robot and their unique ways of interacting with it. In a word, it is necessary to test our designs on diverse tasks in real-world DMS settings with different user populations, to verify the generality of our results and further enrich our findings.

7 CONCLUSION AND FUTURE WORK

In this paper, we systematically formulate three levels of service robot proactivity (*high, medium, low*) based on the degree of anticipation and autonomy, and construct the behavior policy for each level in a decision-making support (DMS) setting. Our experimental results show that users welcome information that can help them narrow down the scope of choices, but dislike unintended intrusion into their decision space. They thus find medium-proactivity robot providing the right level of support in the most acceptable manner. Although users can adjust their behaviors to service robots' way of acting, it is better to have the robots adapt to user needs and contexts. In the future, we will automate robot anticipation by designing a multi-modality model to infer decision-makers' mental state for behavior adaptation, and deploy them in diverse real-world DMS scenarios for field studies.

REFERENCES

[1] Bon Adriel Aseniero, Tiffany Wun, David Ledo, Guenther Ruhe, Anthony Tang, and Sheelagh Carpendale. 2015. STRATOS: Using Visualization to Support Decisions in Strategic Software Release Planning. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1479–1488. https://doi.org/10.1145/2702123.2702426

[2] Jimmy Baraglia, Maya Cakmak, Yukie Nagai, Rajesh Rao, and Minoru Asada. 2016. Initiative in Robot Assistance During Collaborative Task Execution. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI '16)*. IEEE Press, Piscataway, NJ, USA, 67–74. http://dl.acm.org/citation.cfm?id=2906831.2906844

[3] Jenay M. Beer, Arthur D. Fisk, and Wendy A. Rogers. 2014. Toward a Framework for Levels of Robot Autonomy in Human-robot Interaction. *J. Hum.-Robot Interact.* 3, 2 (July 2014), 74–99. https://doi.org/10.5898/JHRI.3.2.Beer

[4] Timothy Bickmore and Justine Cassell. 2001. Relational Agents: A Model and Implementation of Building User Trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. ACM, New York, NY, USA, 396–403. https://doi.org/10.1145/365024.365304

[5] Dan Bohus and Eric Horvitz. 2014. Managing Human-Robot Engagement with Forecasts and... Um... Hesitations. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI '14)*. ACM, New York, NY, USA, 2–9. https://doi.org/10.1145/2663204.2663241

[6] Dan Bohus, Chit W. Saw, and Eric Horvitz. 2014. Directions Robot: In-the-wild Experiences and Lessons Learned. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS '14)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 637–644. http://dl.acm.org/citation.cfm?id=2615731.2615835

[7] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[8] P. Brown, S.C. Levinson, S.C. Levinson, and J.J. Gumperz. 1987. *Politeness: Some Universals in Language Usage*. Cambridge University Press. https://books.google.com.hk/books?id=OG7W8yA2XjcC

[9] T. Chakraborti, S. Kambhampati, M. Scheutz, and Y. Zhang. 2017. AI Challenges in Human-Robot Cognitive Teaming. *ArXiv e-prints* (July 2017). arXiv:cs.AI/1707.04775

[10] Li Chen and Pearl Pu. 2012. Critiquing-based Recommenders: Survey and Emerging Trends. *User Modeling and User-Adapted Interaction* 22, 1-2 (April 2012), 125–150. https://doi.org/10.1007/s11257-011-9108-6

[11] J. Michael Crant. 2000. Proactive Behavior in Organizations. *Journal of Management* 26, 3 (2000), 435–462. https://doi.org/10.1177/014920630002600304

[12] Kerstin Dautenhahn. 2007. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362, 1480 (2007), 679–704. https://doi.org/10.1098/rstb.2006.2004

[13] Julia Fink, Séverin Lemaignan, Pierre Dillenbourg, Philippe Rétornaz, Florian Vaussard, Alain Berthoud, Francesco Mondada, Florian Wille, and Karmen Franinović. 2014. Which Robot Behavior Can Motivate Children to Tidy Up Their Toys?: Design and Evaluation of "Ranger". In *Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction (HRI '14)*. ACM, New York, NY, USA, 439–446. https://doi.org/10.1145/2559636.2559659

[14] Kerstin Fischer. 2011. Interpersonal Variation in Understanding Robots As Social Actors. In *Proceedings of the 6th International Conference on Human-robot Interaction (HRI '11)*. ACM, New York, NY, USA, 53–60. https://doi.org/10.1145/1957656.1957672

[15] The International Organization for Standardization. 2012. Robot Autonomy. https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-2:v1:en.

[16] A. Garrell, M. Villamizar, F. Moreno-Noguer, and A. Sanfeliu. 2013. Proactive behavior of an autonomous mobile robot for human-assisted learning. In *2013 IEEE RO-MAN*. 107–113. https://doi.org/10.1109/ROMAN.2013.6628463

[17] E.R. Girden. 1992. *ANOVA: Repeated measures*. Sage Publications, Inc., Thousand Oaks, CA, US.

[18] Jennifer Goetz and Sara Kiesler. 2002. Cooperation with a Robotic Assistant. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems (CHI EA '02)*. ACM, New York, NY, USA, 578–579. https://doi.org/10.1145/506443.506492

[19] Michael A. Goodrich and Alan C. Schultz. 2007. Human-robot Interaction: A Survey. *Found. Trends Hum.-Comput. Interact.* 1, 3 (Jan. 2007), 203–275. https://doi.org/10.1561/1100000005

[20] Adam M. Grant and Susan J. Ashford. 2008. The dynamics of proactivity at work. *Research in Organizational Behavior* 28 (2008), 3 – 34. https://doi.org/10.1016/j.riob.2008.04.002

[21] Tianran Hu, Anbang Xu, Zhe Liu, Quanzeng You, Yufan Guo, Vibha Sinha, Jiebo Luo, and Rama Akkiraju. 2018. Touch Your Heart: A Tone-aware Chatbot for Customer Care on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 415, 12 pages. https://doi.org/10.1145/3173574.3173989
[22] Chien-Ming Huang, Maya Cakmak, and Bilge Mutlu. 2015. Adaptive Coordination Strategies for Human-Robot Handovers. In *Robotics: Science and Systems*.
[23] Chien-Ming Huang and Bilge Mutlu. 2016. Anticipatory Robot Control for Efficient Human-Robot Collaboration. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI '16)*. IEEE Press, Piscataway, NJ, USA, 83–90. http://dl.acm.org/citation.cfm?id=2906831.2906846
[24] Merkert Johannes, Mueller Marcus, and Hubl Marvin. 2015. A Survey of the Application of Machine Learning in Decision Support Systems. In *ECIS 2015 Completed Research Papers*. https://doi.org/10.18151/7217429
[25] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. 2009. An affective guide robot in a shopping mall. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09*. 173. https://doi.org/10.1145/1514095.1514127
[26] Yusuke Kato, Takayuki Kanda, and Hiroshi Ishiguro. 2015. May I Help You?: Design of Human-like Polite Approaching Behavior. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. ACM, New York, NY, USA, 35–42. https://doi.org/10.1145/2696454.2696463
[27] Peter G.W. Keen and Michael S. Scott Morton. 1978. *Decision support systems: an organizational perspective*. Addison-Wesley.
[28] T. Kim and P. Hinds. 2006. Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction. In *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*. 80–85. https://doi.org/10.1109/ROMAN.2006.314398
[29] Yoshinori Kobayashi, Takashi Shibata, Yosuke Hoshi, Yoshinori Kuno, Mai Okada, and Keiichi Yamazaki. 2010. Selective Function of Speaker Gaze Before and During Questions-Towards Developing Museum Guide Robots. In *CHI EA '10 Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*. 4201–4206. https://doi.org/10.1145/1753846.1754126
[30] H. S. Koppula and A. Saxena. 2016. Anticipating Human Activities Using Object Affordances for Reactive Robotic Response. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, 1 (Jan 2016), 14–29. https://doi.org/10.1109/TPAMI.2015.2430335
[31] Min Kyung Lee, Sara Kielser, Jodi Forlizzi, Siddhartha Srinivasa, and Paul Rybski. 2010. Gracefully Mitigating Breakdowns in Robotic Services. In *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction (HRI '10)*. IEEE Press, Piscataway, NJ, USA, 203–210. http://dl.acm.org/citation.cfm?id=1734454.1734544
[32] Min Kyung Lee, Sara Kiesler, and Jodi Forlizzi. 2010. Receptionist or Information Kiosk: How Do People Talk with a Robot?. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. ACM, New York, NY, USA, 31–40. https://doi.org/10.1145/1718918.1718927
[33] Manja Lohse, Reinier Rothuis, Jorge Gallego-Pérez, Daphne E. Karreman, and Vanessa Evers. 2014. Robot Gestures Make Difficult Tasks Easier: The Impact of Gestures on Perceived Workload and Task Performance. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1459–1466. https://doi.org/10.1145/2556288.2557274
[34] International Federation of Robotics. 2012. Service Robots. https://ifr.org/service-robots/.
[35] Amit Kumar Pandey, Muhammad Ali, and Rachid Alami. 2013. Towards a Task-Aware Proactive Sociable Robot Based on Multi-state Perspective-Taking. *International Journal of Social Robotics* 5, 2 (01 Apr 2013), 215–236. https://doi.org/10.1007/s12369-013-0181-3
[36] Sharon K Parker and Catherine G Collins. 2010. Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management* 36, 3 (2010), 633–662. https://doi.org/10.1177/0149206308321554 arXiv:arXiv:1011.1669v3
[37] Hannah R.M. Pelikan and Mathias Broth. 2016. Why That Nao?: How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4921–4932. https://doi.org/10.1145/2858036.2858478
[38] V. Perera and M. Veloso. 2017. Learning to understand questions on the task history of a service robot. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 304–309. https://doi.org/10.1109/ROMAN.2017.8172318
[39] Pearl Pu, Li Chen, and Rong Hu. 2011. A User-centric Evaluation Framework for Recommender Systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems (RecSys '11)*. ACM, New York, NY, USA, 157–164. https://doi.org/10.1145/2043932.2043962
[40] Pei Luen Patrick Rau, Ye Li, and Jun Liu. 2013. Effects of a social robot's autonomy and group orientation on human decision-making. *Advances in Human-Computer Interaction* 2013 (2013), 6–8. https://doi.org/10.1155/2013/263721
[41] Laurel D. Riek. 2012. Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. *J. Hum.-Robot Interact.* 1, 1 (July 2012), 119–136. https://doi.org/10.5898/JHRI.1.1.Riek
[42] Satoru Satake, Takayuki Kanda, Dylan F. Glas, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. How to Approach Humans?: Strategies for Social Robots to Initiate Interaction. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI '09)*. ACM, New York, NY, USA, 109–116. https://doi.org/10.1145/1514095.1514117
[43] Ameneh Shamekhi, Q. Vera Liao, Dakuo Wang, Rachel K. E. Bellamy, and Thomas Erickson. 2018. Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 391, 13 pages. https://doi.org/10.1145/3173574.3173965
[44] Thomas Sheridan, W L. Verplank, and T L. Brooks. 1978. Human and Computer Control of Undersea Teleoperators. *Advances in Human-Computer Interaction* (1978), MIT Man–Machine Systems Laboratory.
[45] Chao Shi, Michihiro Shimada, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2011. Spatial Formation Model for Initiating Conversation. In *Proceedings of Robotics: Science and Systems*. Los Angeles, CA, USA. https://doi.org/10.15607/RSS.2011.VII.039
[46] Masahiro Shiomi, Kazuhiko Shinozawa, Yoshifumi Nakagawa, Takahiro Miyashita, Toshio Sakamoto, Toshimitsu Terakubo, Hiroshi Ishiguro, and Norihiro Hagita. 2013. Recommendation Effects of a Social Robot for Advertisement-Use Context in a Shopping Mall. *International Journal of Social Robotics* 5, 2 (01 Apr 2013), 251–262. https://doi.org/10.1007/s12369-013-0180-4
[47] Jacob Solomon. 2014. Customization Bias in Decision Support Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3065–3074. https://doi.org/10.1145/2556288.2557211
[48] Vasant Srinivasan and Leila Takayama. 2016. Help Me Please: Robot Politeness Strategies for Soliciting Help From Humans. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4945–4955. https://doi.org/10.

1145/2858036.2858217
[49] Mingfei Sun, Zhenjie Zhao, and Xiaojuan Ma. 2017. Sensing and Handling Engagement Dynamics in Human-Robot Interaction Involving Peripheral Computing Devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 556–567. https://doi.org/10.1145/3025453.3025469
[50] Leila Takayama, Victoria Groom, and Clifford Nass. 2009. I'm Sorry, Dave: I'm Afraid I Won'T Do That: Social Aspects of Human-agent Conflict. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 2099–2108. https://doi.org/10.1145/1518701.1519021
[51] Cristen Torrey, Susan Fussell, and Sara Kiesler. 2013. How a Robot Should Give Advice. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction (HRI '13)*. IEEE Press, Piscataway, NJ, USA, 275–282. http://dl.acm.org/citation.cfm?id=2447556.2447666
[52] L. Woiceshyn, Y. Wang, G. Nejat, and B. Benhabib. 2017. Personalized clothing recommendation by a social robot. In *2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*. 179–185. https://doi.org/10.1109/IRIS.2017.8250118
[53] Yanzhen Yue, Xiaojuan Ma, and Zhenhui Jiang. 2014. Share Your View: Impact of Co-navigation Support and Status Composition in Collaborative Online Shopping. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3299–3308. https://doi.org/10.1145/2556288.2557143
[54] Yanzhen Yue, Xiaojuan Ma, and Zhenhui Jiang. 2016. Influence of Content Layout and Motivation on Users' Herd Behavior in Social Discovery. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5715–5719. https://doi.org/10.1145/2858036.2858497
[55] Yu Zhang, V. Narayanan, T. Chakraborti, and S. Kambhampati. 2015. A human factors analysis of proactive support in human-robot teaming. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 3586–3593. https://doi.org/10.1109/IROS.2015.7353878