

Learning-Enabled Robotic Assistive Support: Understanding Older Adult Opinions and Comparing Them to Younger Adult Opinions

Nisha Raghunath, M.S.

Washington State University Department of Psychology, Pullman, USA

Christopher Pereyda

Washington State University School of EECS, Pullman, USA

Justin F. Frow, B.S.

Washington State University Department of Psychology, Pullman, USA

Diane Cook, Ph.D.

Washington State University School of EECS, Pullman, USA

Maureen Schmitter-Edgecombe, Ph.D.

Washington State University Department of Psychology, Pullman, USA

Corresponding Author: Maureen Schmitter-Edgecombe, Department of Psychology, Washington State University, P.O. Box 644820, Pullman, Washington 99164-4820, Phone: 509-335-0170, Fax: 509-335-5043, e-mail: schmitter-e@wsu.edu

Acknowledgments:

This work was partly supported by Grants No. 1734558 and 1543656 from the National Science Foundation. The authors would like to thank Shivam Goel, Sepehr Nesaei, Gabriel de la Cruz, Garrett Wilson and Matthew Taylor for their contributions to this research project.

Abstract

Background

As older adults age, they may require assistance completing activities of daily living (ADLs). Robotic assistance can offset healthcare costs and allow older adults to preserve their autonomy. Younger adults are often involved in the design and purchase of these robotic technologies, and must take into account the needs and expectations of the target population (i.e., older adults) to create a robot that the end-user will adopt.

Research Aim/Questions

This study evaluated the opinions of both younger and older adults regarding the design and performance of the Robot Activity Support (RAS) system. It is important to understand points of agreement and divergence between these populations' perspectives so that effective robotic aids are created for older adults.

Methods

Fifty-two younger and older adults completed three scripted tasks with the RAS robot in a smart home environment. Each participant made task errors to cue the robot to offer help via three prompt modalities (guide to the object, video of forgotten step, and video of the full task). After interacting with the robot, participants completed questionnaires to evaluate their opinions of and satisfaction with the robot.

Results

There were no differences between younger and older adults' perceptions of the robot across a variety of factors (e.g., likability, cognitive demand), with both age groups expressing generally neutral opinions. Both groups rated the Full Video prompt as least helpful, effective, and liked. Participants recommended the robot's response accuracy, movement speed, alerting style and system flexibility be improved. Younger adults overestimated how much older adults would want a robot like this.

Conclusions

This study underscores the importance of testing technology with the end-user, as older adults were less interested in having a similar robot in their home than younger counterparts expected. Future studies will show if older adults' opinions can be improved after making the recommended changes.

Keywords: robot, autonomous, assistance, mild cognitive impairment, activities of daily living, smart home

INTRODUCTION

As older adults (OAs) age, they are more likely to experience chronic diseases and cognitive decline, and may begin to need assistance completing activities of daily living (ADLs) such as medication adherence and cooking. Due to the increasing costs of health care and the insufficient number of available caregivers (both family and professional), utilizing technology to supplement caregiving is an important goal. Technology use may also enable OAs to age in place and maintain more autonomy. Robotic technology has gained momentum as a means to assist OAs in their everyday lives. For example, several robots such as Guido and RIBA-II are built to help with movement (Martins, Santos, Frizera & Ceres, 2015; Jiang, Huang, Huo, Zhang, & Song, 2018). Paro and Care-O-bot 4 are examples of service and companionship robots developed to encourage interaction with the robot and promote social skills and cognitive well-being (Yu et al., 2015; Kittmann et al., 2015). Other robots include Ed, who is teleoperated to provide audio/video step-by-step help to OAs during a task, and PEARL, which reminds individuals to initiate activities of daily living (e.g., meals and daily hygiene; Wang, Sudhama, Begum, Huq, & Mihailidis, 2017; Pollack et al., 2002).

Memory impairment, the hallmark symptom of mild cognitive impairment (MCI), can interfere with the successful completion of ADLs. All these robotic aids, while designed to support OAs in their daily routines, lack one thing: they are not capable of helping OAs when they are unable to remember how to complete a task without human intervention. To this end, we created the Robot Activity Support (RAS) system and robot (Wilson et al., 2019). Paired with a smart home environment, the RAS robot is able to monitor a resident's activity patterns, recognize when an error is made in task execution and intervene to offer help in a multi-modal fashion.



The robot (RAS) is an amalgamation of elements, each with a specific purpose (see Figure 1). At the top of the robot is a RGBD camera that allows the robot to detect both the objects involved in everyday tasks, and the human in its environment (Wilson et al., 2019). Located at the bottom of the robot are the components of the navigation manager (e.g., the LiDAR and computational units) and the hardware required for the robot to maneuver in a space (e.g., wheels and battery; Wilson et al., 2019). The robot also has a mounted Android tablet, which serves as the point of interaction for users. When users forget to perform a step in a task, the missing step is detected by the RAS system due to a

lack of interaction with the forgotten object's attached accelerometer-based motion sensor

Figure 1. RAS robot with human detection, user-interface and navigation components.

(Estimote; Estimote, Inc., 2012). Immediately upon error detection, the tablet plays an audio clip asking the participant if they need help (i.e., "Hello, how can I help you?") and begins approaching the person. The screen display changes from a neutral facial expression to a surprised one and displays the question "Do you want help?" with a Yes and No button situated below the question (see Figure 2b). If the participant selects "Yes", the tablet displays three help prompt choices: (1) a video of the entire task without errors, (2) a video of only the forgotten step, and (3)

an offer to lead the user to the forgotten object. Once finished using the chosen help prompt, the user taps the "I did it!" button and the robot returns to a home position (see Figure 2 for the entire sequence of tablet screen displays). Of importance, all of the robot's capabilities (i.e., task monitoring, error detection, space mapping, robot navigation, object and human finding and the interface) are completely autonomous. That is, human intervention was not necessary for the robot to intervene when necessary and perform the subsequent actions, unlike the work done before RAS.

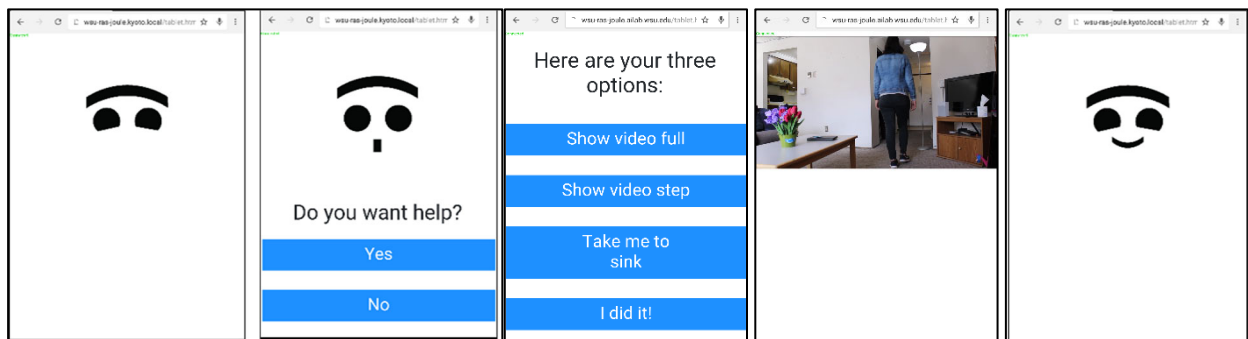


Figure 2. (a) Neutral state face, (b) query face and prompt asking if user needs help, (c) no face while showing help prompt options, (d) video that plays when either of the video help prompts is selected, and (e) error-corrected face as robot retreats.

Though this robot is designed for OAs' use, especially those with memory impairment, the younger generation is typically involved in the design and selection of these technologies. Thus, it is important to understand both similarities and differences in older and younger adults' attitudes toward and expectations of robotic assistive aids. Previous literature informs us that younger adults prefer a robot with more human-like characteristics while older adults prefer a more anthropomorphic look, perhaps due to raised expectations of capabilities that come with a humanoid appearance that the robot cannot fulfill (Pino, Boulay, Jouen, & Rigaud, 2015; Stafford, MacDonald, Li, & Broadbent, 2014). Older adults also report preferring a female voice/appearance, slower speed and smaller (i.e., not human-like) size for robots (Stafford, MacDonald, Li, & Broadbent, 2014).

Appearance-opinions aside, the biggest factor in acceptance of robotic assistance that emerged from past studies was an individual's need for it. Younger adults were more likely to think of robots as devices that perform boring, household chores (e.g., cleaning), and were amenable in general to including a robot in their everyday life (Goher, Mansouri, & Fadlallah, 2017; Smarr et al., 2015). Older adults also considered robots useful but for different purposes, and only when the robots enabled them to maintain their independence (Wu et al., 2016). Older adults appear most interested in robots that have the potential to effectively and efficiently help them complete everyday tasks and ensure safety by detecting falls and issuing danger warnings (Smarr et al., 2015). Furthermore, the robot's ability to perform tasks efficiently and with minimum human-intervention seemed to most strongly determine acceptance from both groups (Smarr et al., 2015).

Differences between younger and older adults' expectations and preferences about robotic assistance can inform the design of robotic devices, as those who create robotic technology and play a role in technology selection are often younger adults. If the design and capabilities of robotic devices do not align with older adults' expectations, this will negatively affect OAs' acceptance of and willingness to use the assistive technology. Accordingly, in this study, we evaluate and compare YA and OA perceptions of our RAS robot designed to help complete ADLs.

METHODS

Participants

Fifty-two individuals, 26 younger adults (ages 18-29) and 26 older adults (ages 52-87), represented the final sample for this RAS usability testing study. The data from four younger adults and one older adult were discarded from the original sample due to robot malfunctions ($N = 2$), age outside the range ($N = 2$), and incomplete or incorrectly-completed questionnaires ($N = 1$). Table 1 provides a summary of participant demographics. Fifty percent of each group ($N = 13$) were female. As expected, older adult participants were significantly older than younger adult participants, $t(48) = 22.96, p < .001$, and had significantly more years of education, $t(50) = 6.29, p < .001$. Among the OAs, 34.62% of participants fell in the range of mild cognitive impairment on a brief screening measure of cognitive abilities (i.e., the Montreal Cognitive Assessment; *MoCA*; Nasreddine et al., 2005). The *MoCA* requires participants to complete brief cognitive tasks such as backward digit span (i.e., repeat a sequence of numbers in the reverse order) and language fluency (i.e., name as many words as possible that begin with the letter F in one minute) tasks. A score of 26 or higher out of 30 is considered to be within normal limits.

Younger adult participants were recruited via Washington State University's undergraduate psychology experiment recruitment site (SONA) or from Washington State University's Research Experience for Undergraduates Summer Program (REU). Older adult participants were recruited from prior studies conducted through Washington State University's Neuropsychology and Aging laboratory. This study was approved by the University Institutional Review Board.

Table 1. Participant demographics (mean and standard deviations) of younger and older adults.

	Young Adults	Older Adults
Participants	N = 26	N = 26
Age (yrs.)	23.00 (2.85)	71.00 (10.12)*
Gender (% females)	50%	50%
Education (yrs.)	14.00 (1.10)	17.00 (1.70)*
MoCA	---	26.50 (2.89)
Comfort Using Technology	6.12 (0.89)	5.80 (0.92) ^a

Notes. MoCA = Montreal Cognitive Assessment. ^aRepresents N = 24. * $p < .001$

Assessment Instruments

The following instruments were used to gather information about participant comfort with technology, evaluation of robot characteristics, and self-report of satisfaction and usability of the robot.

Technology Comfort Questionnaire (TCQ); derived from Tam, Van Son, Dyck & Schmitter-Edgecombe, 2017). To understand individuals' comfort with technology, we asked participants to rate six statements, such as "I am generally comfortable with technology," on a Likert scale of 1 (Strongly Disagree) to 7 (Strongly Agree). Scores were calculated by averaging all responses for an overall score. Higher scores indicate greater comfort with using technology (Cronbach's $\alpha = .68$).

Subjective Assessment of Speech System Interfaces (SASSI); Hone & Graham, 2000). The SASSI is comprised of 34 statements, rated on a scale of 1 (Strongly Disagree) to 7 (Strongly Agree), that assesses a users' subjective evaluation of a speech input/output system to predict success of the system for everyday use. The questionnaire includes the following six subscales that factor into users' perceptions: system response accuracy, likeability, cognitive demand, annoyance, habitability and speed. Reliability coefficient alphas of .90, .91, .88, .77, .75 and .69, respectively, demonstrate acceptable to excellent validity for the first five factors and borderline acceptable validity for the sixth factor (Hone & Graham, 2000). The subscale scores were calculated by reverse-scoring the appropriate items and then averaging responses for each subscale; there is no overall score.

Robot Characteristics Questionnaires (RCQ). These surveys were generated by the research team to assess participants' opinions about the robot's design, namely its movement and prompting abilities. Participants provided Likert ratings about several aspects of RAS (e.g., height, speed) and the help prompts (i.e., Guide to Object, Step Video, Full Video) on a scale of 1 (Strongly Disagree) to 5 (Strongly Agree). Each help prompt was assessed for four separate purposes, namely ease of imitation, helpful if unable to recall next step, confusing to someone with MCI, and helpful to someone with MCI. Participants also indicated which of the three help prompts they best liked, found most helpful and deemed least effective. Open-ended questions

further provided participants with an opportunity to offer their opinions about what they liked and did not like about the robot. Participants were also asked for what purposes an older adult (they themselves) or someone with MCI might find the robot useful. Other questions asked participants for suggestions for helpful alternative prompting mechanisms and general improvements that could be made to the robot.

Procedure

For the duration of the interactive testing sessions with the robot, two experimenters were stationed upstairs in a control room and monitored the participant via the smart apartment's webcams. Thus, the participant was the only human with RAS in the testing area on the ground level of the apartment. One of the experimenters monitored RAS' program and navigation terminals to assist when necessary with manual fine tuning of the robot's movement. The second experimenter communicated with participants via intercom, leading participants through each scripted step of the three tasks detailed below. This experimenter also monitored accuracy of data collection from the Estimote sensors, which were attached to objects of interest being used in the three activities.

When participants arrived at the campus smart apartment testbed, they were brought to the living room where the second experimenter obtained informed consent. After this, the experimenter explained the study's purpose and how the experiment would proceed, the participant was given a tour of the apartment's lower level to orient them to where each task-relevant object was located. Once finished, the experimenter went upstairs to the control room, from where they led the participant through the study tasks and interaction with the robot.

Representative of everyday daily living activities, the tasks in this study were (1) preparing to walk the dog, (2) taking medication with food and water, and (3) watering the apartment plants. Participants completed each task once in the correct sequence of event steps without mistakes, and three additional times, each time with a different error. Regarding these errors, for the preparing to walk the dog task, participants skipped retrieving the umbrella, then the dog leash, then the apartment keys, interacting with the robot after each error in three separate task runs. For the take medication with food and water task, participants neglected to retrieve a granola bar, followed by the medication bottle, and lastly failed to take medication out of the bottle. For the watering apartment plants task, participants forgot to fill the watering can with water, then water the coffee table plant, then water the side table plant. For a full description of the exact sequence of each task and its corresponding errors, see Wilson et al., (2019). Each participant completed all the tasks in the exact same order, and the errors of omission were not varied across participants. The type of help prompt chosen (i.e., Guide to Object, Step Video and Full Video), however, was counterbalanced across task errors for each group. As participants completed these tasks and interacted with the robot and tablet interface, the second experimenter made note of adverse events that occurred (i.e., when the robot did not operate as expected or the participant experienced difficulties trying to interact with it).

Upon detecting an error, the robot audibly asked if the participant needed help and began approaching them. Once the robot came to a stop within reaching distance, the person was

instructed to accept the offered help by pressing “Yes” and then choose the prompt type that was designated for that specific trial. After selecting the prompt as directed by the experimenter and either watching a video or following the robot to an object, participants completed the forgotten step as directed. Following completion of the error step, participants then tapped the “I did it!” button, effectively dismissing the robot; participants then finished the remainder of the task without additional errors.

Once finished with the robot interaction portion of the experiment, participants completed a series of questionnaires compiled either in the online survey platform Qualtrics (Qualtrics, Provo, UT) or in paper-and-pen format. The entire study lasted approximately two hours. Younger adult participants were awarded 2.0 credit hours for their psychology class research quota through the SONA system, or recorded 2.0 hours of work for the REU program. Older adult participants were given a \$25 cash honorarium.

RESULTS

Analysis

We begin by reporting error rate data about the robot performance. Next, using an independent samples t-test, the data was examined for differences between groups (YA and OA) on the questionnaire examining comfort using technology. To assess for group differences in RAS ratings and opinions, the six SASSI questionnaire subscale scores were compared between groups using repeated measures analyses of variance with Bonferroni post-hoc tests. Responses to the RCQ Likert questions were compared between groups using independent samples t-tests. Responses to open-ended qualitative questions were examined for commonalities among participants. We also conducted repeated measure analyses of variance with prompt type as the within-subject variable, and used Bonferroni post-hoc tests to examine differences across prompt types. Opinions of each help prompt (Guide to Object, Step Video, and Full Video) were analyzed using repeated chi-square tests.

Robot Performance (Error Rates)

Using experimenter notes from the testing sessions, we tabulated the number of successful task runs (i.e., tasks completed correctly by the robot without operator intervention). Each testing session consisted of 12 tasks, 9 of which expected the robot to intervene and offer help (the remaining 3 tasks were scripted to occur correctly without any errors for the robot to detect), for a total of 624 tasks across all 52 YA and OA participants. However, since some tasks were done multiple times when errors occurred to allow participants to experience each prompt type with each task while others were not performed due to time constraints, 683 total tasks were completed. Across all testing sessions, 371 tasks were completed correctly (i.e., the robot prompted and moved properly at the appropriate time) without operator intervention for a task success rate of 0.544. Errors that occurred were due to robot error (i.e., robot did not face human or remained more than 2 feet away from the human/object, robot became stuck on object in apartment or stopped moving to human/object), interface error (i.e., tablet did not register user

response, options did not display correctly, tablet did not play verbal prompt, tablet provided prompt without system initiation), system error (i.e., system froze, communication was lost between Estimote and server or server and robot), or Estimote error [false negatives (i.e., the participant made an error that the system/interface did not detect), and false positives (i.e., the participant did not appear to make an error but the system/interface still prompted them)].

Technology Comfort Level

As seen in Table 1, the older and younger groups did not differ in their reported level of comfort using technology, $t(44) = -1.26, p = .21$, which was high for both groups (≥ 5.8 out of 7.0).

SASSI

A repeated measure analysis of variance was conducted to determine whether SASSI subscales differed between groups and from one another. The analysis revealed no overall differences between the younger ($M = 4.74$) and older ($M = 4.42$) adult groups in their average response ratings to questions, $F(1,48) = 1.49, p = .23$, which fell within the Neutral to Slightly Agree range (see Table 2). There was, however, a main effect of subscales, $F(1,48) = 49.82, p < .001$. Post hoc analyses revealed that Cognitive Demand of the RAS system was the highest rated subscale ($M = 5.48$) and this subscale differed significantly from all other subscales with the exception of Likeability, which was the second highest rated subscale ($M = 5.31$). Habitability was the third highest rated subscale ($M = 4.97$) and differed significantly from all subscales but Likeability. System Response Accuracy was rated fourth ($M = 4.58$) and differed significantly from the Annoyance subscale ($M = 4.12$). Furthermore, the Annoyance subscale differed significantly from the lowest rated Speed subscale ($M = 3.11$), which fell into the Slightly Disagree range. Importantly, there was no interaction between group and subscales, $F(1,48) = 1.02, p = .41$, suggesting that the pattern of responses across subscales was similar for the older and younger adult groups.

Table 2. *Subjective Assessment of Speech System Interface subscale scores for both younger and older adult groups.*

	YA Average (SD)	OA Average ^a (SD)
<u>SASSI</u>		
System Response Accuracy	4.93 (1.24)	4.23 (0.25) ^a
Likeability	5.45 (1.29)	5.16 (0.24) ^a
Cognitive Demand	5.72 (0.85)	5.23 (0.18) ^a
Annoyance	4.19 (1.18)	4.12 (0.16) ^a
Habitability	4.95 (1.03)	4.97 (0.16) ^a
Speed	3.19 (1.77)	3.02 (0.24) ^a

Notes: ^aAnalyses conducted with 24 older adult participants; SASSI = Subjective Assessment of Speech System Interface.

Robot Characteristics Questionnaire

Participants' responses to the first three questions on the RCQ (i.e., the robot would successfully get my attention if it moved and worked like this, the robot would startle me if it moved like this, and the height of the robot does not intimidate me or make me feel uncomfortable) were compared using independent samples t-tests with age group as the between-subjects factor. Answers to the second question were reverse-scored. Analyses showed that older and younger adults did not significantly differ in their opinions, rating each query in the Neutral to Agree range [successfully get my attention = 1.00 – 5.00 (OA and YA); would startle me = 3.00 – 5.00 (OA) and 1.00 – 5.00 (YA); height is not intimidating = 2.00 – 5.00 (OA) and 1.00 – 5.00 (YA)], $t_s(50) > -0.25$, $p_s > .20$.

When asked what distance between a user and the robot would be most comfortable, older and younger adults again expressed no significant differences with an average answer of 21.55 inches from younger adults and 23.53 inches from older adults, $t(37) = -.51$, $p = .11$. Nine participants (3 OA and 6 YA) did not report a numeric distance but a qualitative answer. Three expressed preferring the robot to be farther away, four reported wanting it to come closer, and two were not specific in their preferences (e.g., "... comfortable with any distance" and "A typical distance between two friends").

After observing the robot's performance and taking note of participants' verbal opinions, we opted to additionally survey OA participants about their willingness to wait for the robot should it need to move slowly, how often they would tolerate false information from the robot and how often they would tolerate the robot approaching when unwanted. Older adult participants rated these statements on a scale of 1 (Never) to 5 (Always). On average, OAs were relatively unwilling to wait and intolerant of the robot's errors; they indicated being willing to wait sometimes to about half the time ($M = 2.75$, $n = 24$), being tolerant of the robot providing incorrect information sometimes ($M = 2.00$, $n = 24$) and being tolerant of the robot approaching when help is not needed sometimes to about half the time ($M = 2.5$, $n = 24$).

To assess whether participants would desire to adopt our robot, we asked younger adults whether they believed their older relatives would enjoy having our robot in their home, and asked older adults if they would enjoy having our robot in their home. Younger adults' responses fell in the Neutral to Agree range on a Likert scale of 1 to 5 ($M = 3.59$) while older adults disagreed with this statement ($M = 2.59$), and this difference in opinions was significant, $t(47) = -3.26$, $p = .002$.

Free Response Open-Ended RCQ Items

Participants' responses to the open-ended questions of the RCQ were analyzed for common themes. The most common responses related to things participants reported liking about the robot included its simple and accessible design (e.g., "simple to use", "simple seeming") and how quickly it recognized errors (e.g., "alerted me right away that I missed a step in the activity"). Participants indicated that they did not like the robot's speed/movement (e.g., too slow

in coming to help me with task”, “jerky travel”, “it moves slowly and I got impatient”), and were made uncomfortable by some aspects of the robot, including its “awkward movements” and possible “tripping hazard”. The most common responses about the types of tasks the robot would be most useful for included reminders (e.g., to take medication, phone calls, objects’ locations) and assistance completing tasks (e.g., making meals, animal care). For individuals with memory impairment, participants thought the robot would be most useful for tasks related to medication and hygiene, or more generally to “tasks most important to their everyday health”. Participants also thought the robot would be helpful for breaking down everyday tasks into individual steps and guiding participants through the steps.

Regarding general suggestions for improvement of the robot, common themes emerged included wanting faster/smoothier movements, more alerts, better hardware (i.e., a “nicer” robot that looks “a bit more friendly”) and higher robot function reliability (e.g., “system had a few false positives”). For example, participants expressed wanting the robot “to be able to move beyond monitoring a strict chain of events”, and would like it better “if steps could be done in any order.” They also recommended “adding sensor accuracy, computing power” and an “easier or more responsive tablet.” Furthermore, participants recommended adding some sort of cue to get the users attention before the robot started moving; suggestions included add an “alert system – music or flashing light”, “beeper and flashing LED on objects”, and “maybe a blinker or musical tone.”

Prompt Ratings Related to Specific Purposes

Separate repeated measure analyses of variance were conducted for each of the four prompt purpose questions (see Table 3). Answers to the third purpose (confusing to someone with MCI) were reverse-scored so they were in the same direction as the other items. Analyses revealed no significant main effect of age group or interactions between age group and prompt type for any of the four prompt purpose questions. There was a main effect of prompt type when rating each prompt type for how helpful it would be if unable to recall the next step in a task, $F(2,84) = 4.97, p = .009$, and for someone with MCI, $F(2,84) = 6.57, p = .004$. Post-hoc tests revealed that the Full Video was rated as less helpful than the Next Step Video and Guide to Object for both the questions evaluating helpfulness related to next step in task ($ts > 2.17, ps < .05$) and for someone with MCI ($ts > 2.44, ps < .05$). There were no differences in ratings between the three prompt types for questions evaluating ease of imitation, $F(2,84) = 2.59, p = .08$, or how confusing the prompt types might be to someone with MCI, $F(2,84) = 2.75, p = .07$.

Table 3. *Participant ratings (mean and standard deviation values) of each prompt type for four separate purposes.*

Prompt Purpose	Prompt Type		
	Guide to Object	Next Step Video	Full Video

Age Group	YA	OA	YA	OA	YA	OA
Ease of Imitation	3.83 (0.94)	4.05 (1.07)	4.00 (0.85)	4.10 (0.89)	3.52 (1.12)	3.86 (1.28)
Helpful if Unable to Recall Next Step	4.04 (0.78)	4.00 (0.95)	4.22 (0.67)	4.05 (1.16)	3.57 ^{a,b} (1.20)	3.67 ^{a,b} (1.39)
Confusing to Someone with MCI	3.38 (1.10)	3.29 (1.00)	3.79 (0.93)	3.30 (1.06)	3.26 (1.05)	3.05 (1.21)
Helpful to Someone with MCI	4.00 (0.80)	3.71 (1.01)	4.04 (0.56)	3.67 (1.11)	3.43 ^{a,b} (1.12)	3.43 ^{a,b} (1.25)

Notes: Analyses conducted with N = 21 OA and 23 YA; ^aSignificantly different value from Guide to Object rating; ^bSignificantly different value from Next Step Video rating.

Preferences between Help Prompts

Chi-square tests of independence revealed significant differences between prompt preferences for each prompt rating category among OAs, $\chi^2(4, N = 65) = 18.91, p = .001$, and YAs, $\chi^2(4, N = 75) = 54.82, p < .001$. For the best liked prompt, post-hoc analyses showed that OAs chose the Full Video prompt (9% of OAs) significantly less than the expected frequency, $\chi^2(4, N = 65) = 5.76, p = .02$. Younger adults chose the Guide to Object, $\chi^2(4, N = 75) = 4.41, p = .04$, and Next Step Video, $\chi^2(4, N = 75) = 13.69, p < .001$, prompts significantly more than the expected frequencies as the best liked prompts. Regarding preferences for the most helpful prompt, the Next Step Video prompt was chosen significantly more than the expected frequencies for both OAs (68% of OAs), $\chi^2(4, N = 65) = 4.41, p = .04$, and YAs, $\chi^2(4, N = 75) = 4.41, p = .04$. YAs further chose the Full Video prompt less than the expected count, $\chi^2(4, N = 75) = 13.69, p < .001$, as the most helpful prompt. Older adults chose the Next Step Video prompt significantly less than the expected frequency as the least effective prompt (a good perception), $\chi^2(4, N = 65) = 11.56, p < .001$, as did younger adults, $\chi^2(4, N = 75) = 18.49, p < .001$. Additionally, YAs also chose the Guide to Object, $\chi^2(4, N = 75) = 6.25, p = .01$, significantly less than the expected frequencies as the least effective prompt (again, a positive perception). Furthermore, OAs chose the Full Video prompt (59% of OAs) significantly more than the expected frequency as the least effective prompt, $\chi^2(4, N = 65) = 16.00, p < .001$, as did YAs, $\chi^2(4, N = 75) = 54.76, p = .04$.

Table 4. *Frequency counts of participants' ratings for most liked, most helpful and least effective prompt styles.*

Prompt Rating Category	Prompt Type		
	Guide to Object	Next Step Video	Full Video

Age Group	YA	OA	YA	OA	YA	OA
Best Liked Prompt	38%	32%	62%	59%	0%	9%
Most Helpful Prompt	38%	18%	62%	68%	0%	14%
Least Effective Prompt	12%	23%	8%	18%	80%	59%

Notes: N =22 OA and 25 YA.

Qualitative Evaluation of Help Prompts

When asked about alternative prompts and changes to said prompts, common themes included making the system voice-interactive and supplementing the motions/videos with lights or images and signs (i.e., arrows and photos of the target forgotten object). Beeps, tones and musical notes were also suggested as useful additions. More than any other suggestion though, participants from both age groups recommended adding auditory instructions (23 participants), and 4 participants specifically said to change the verbiage of the audio prompts. Other common suggestions included adding affirmations when part of the task was completed successfully, and an audio message after the help prompt was complete to ask the user whether they would like to continue receiving help or resume the task. Several individuals commented that the video editing for the Full Video prompt should be smoother, that it was too long, and that it might be beneficial for the robot to ask users “Would you like to see the whole process?” before playing the video.

DISCUSSION

Our goal was to evaluate opinions and perceptions of our robotic activity support system and to assess for differences between younger and older adults’ perceptions, expectations and preferences regarding robotic assistance. We found that participants in both groups expressed generally neutral opinions regarding various aspects of the robot design. Younger and older participants alike slightly agreed that the robot and interface was not mentally taxing to use, but were not satisfied (slightly disagreed) with the speed of the robot. Open-ended comments further supported the need to improve the robots speed, as suggestions for faster/smooth movements emerged as a common general theme for robot improvement. Regarding the help prompts, the Next Step Video and Guide to Object prompts were endorsed by both groups as being significantly more helpful than the Full Video prompt when someone was unable to recall the next step in a task and for someone with MCI.

When asked which of the three prompts they liked best, none of the younger adults and only two of the older adults endorsed liking the Full Video prompt the best. Both age groups found the Next Step Video to be the most helpful, and appeared to like it best, though statistical

analyses did not show this preference to be significant. The Full Video prompt was also rated as the least effective by both age groups. This may partly reflect the added prospective (e.g., remembering to carry out the correct needed step after prompt finished) and retrospective (e.g., watching already completed task steps) memory demands of the Full Video prompt, which could further confuse someone with MCI. Research has shown that individuals with MCI and AD experience both prospective memory (i.e., ability to remember to perform actions at a later time) and retrospective memory difficulties (Dermody, Hornberger, Piguet, Hodges, & Irish, 2015; Hernandez Cardenache, Burguera, Acevedo, Curiel, & Loewenstein, 2014; Hsu, Huang, Tu, & Hua, 2015). In addition, the National Institute on Aging recommends caregivers providing “simple, step-by-step instructions” to those with AD to improve communication (Alzheimer’s Caregiving: Changes in Communication Skills, n.d.). Consistent with this recommendation, the Next Step Video prompt offered small but relevant, simple, and to-the-point aid.

Our original motivation to create the Guide to Object prompt was to enable our robotic assistance to autonomously maneuver to the user/desired items – a feature that set our robot above existing assistive technology. We found that 38% of YAs and 32% of OAs rated Guide to Object as the best liked prompt, and 38% of YAs and 18% of OAs rated it as the most helpful prompt. It is possible that these numbers would have been higher if we had specifically asked participants about a situation where they could not locate an item. It may be that preferences for the best prompt type will change depending on the situation (e.g., difficulty locating item versus difficulty remembering the next step). We also found that a few older adults did rate the Full Step video as best liked and most helpful, indicating that prompt type preferences also vary across individuals. This suggests that some flexibility should be built into the system.

Although participants reported clear preferences regarding the help prompts, overall ratings of the robot were generally neutral. Users’ perceptions of aspects of the robot system (i.e., SASSI ratings) fell within the Neutral to Slightly Agree range (i.e., 4.73 for YAs and 4.46 for OAs out of 7). This indicates that we are currently falling short in the design of our robot interface. In its current form, younger and older adults are rating the robot similarly and are consistent in their ratings of varying factors that impact user perception. Specifically, both age groups rated most positively (in the Slightly Agree to Agree range) aspects of the robot related to cognitive demand and likeability. The cognitive demand subscale assesses the perceived amount of effort required to use the robot and how it made participants feel, and the likeability subscale measured participants’ affective feelings about the system. Ratings between Neutral and Slightly agree were found for characteristics related to annoyance, habitability and system response accuracy. Both groups also rated the speed subscale the lowest (in the Slightly Disagree). The speed subscale measures participants’ opinions about the system’s performance speed. These ratings indicate several characteristics of the robotic support system that especially need to be approved upon. In open-ended questions, participants also suggested that the robot’s movement needs to be smoother (e.g., “less jerky”), and this is consistent with the robotic-interaction literature (Kashi & Levy-Tzedek, 2018).

Regardless of whether we are able to improve our robot such that participants' ratings increase to the upper-most ranges of the SASSI (i.e., 6 to 7 range), we recognize that it is possible to create a highly satisfactory and well-liked robot that individuals do not find to be of value or wish to use. In our case, it is important that our target population (older adults) be interested in using the robot. Younger adults significantly overestimated how much older adults would want the robot in their homes. Numerous older adults reported that although they could see the potential of this technology for being helpful, they did not want or need it at their current stage of life. It is therefore possible that a larger number of OAs may have expressed more desire to adopt the robot had we asked if they would want it *eventually*, at a different stage in their life if they require this type of assistance. As suggested by the open-ended responses, participants conveyed that they appreciated the simplicity of the interface and were impressed with its capabilities and prompt responding when it worked. However, they recommended that future work improve the robot's speed and movement (i.e., faster and smoother, stay out of the way during and after guidance to objects), performance (i.e., fewer false negatives and positives for error detection), and physical design (i.e., friendlier expressions, sturdier build). Given these results, it would appear that our target population is not interested in our robot as it is at this point in their lives, further underscoring the need to improve its features and design to appeal to the older adult population, versus only the younger adults.

Another issue participants had with the robot was the lack of flexibility with which they could complete the task steps. For purposes of our study, the tasks participants completed were scripted and required task steps be performed in a specific order so the system could register each step as correctly completed. As experimenters, we utilized cameras and the Estimotes on the target objects (e.g., umbrella, medication bottle, watering can) to detect completed steps and errors, as well as manually started Estimote monitoring before each task so the robot would sense when a step was skipped. All of these elements are not easily translatable to real-world use, nor were they appreciated by participants. For example, the older adults in particular reported frustration with having to follow directions instead of having the freedom to perform steps as desired. Therefore, for our robot to effectively help users, the robot will need to be capable of tracking activity patterns, storing the step information, and understanding that activities could be completed in varying ways.

There are limitations of our study that prevent us from effectively generalizing our results. Due to the Estimotes' and network connectivity issues in the smart apartment, both of which contributed to the robot's inconsistent performance and accuracy when helping users, perceptions of its capabilities may have been reduced, which could in turn have negatively affected OA's desire to use the robot. However, while a standalone robot may not pose these issues, partnering a robot with other technologies extends its capabilities and therefore shows that we should continue to improve our model and system. In our study set-up, we had to place multiple Estimotes (object sensors) on the target objects used in order for the RAS system to correctly identify completed and missed task steps. Because these have a limited battery life though, it can prove to be a recurring added expense for consumers, necessitating that we

discover an alternate type of object sensor or find a way to eliminate the need for object sensors in future studies. Our study design also necessitated that participants complete the scripted tasks in a specific and inflexible order, undermining the naturalistic element and resulting in lesser opinions of the robot (as noted by participants). Improvement of the robot's speed, accuracy and consistency, as well as adaptation of the robot's software to recognize task steps in a variable order may also lead to higher ratings of the robot.

CONCLUSIONS

Going forward, it is important to consider how we can improve the robot to increase its desirability amongst our target population. Research by Feingold-Polak et al. shows that customizable robots are better rated and enjoyed by older adults (2018), and this is an avenue we should consider for our robot. Multiple OA participants also suggested incorporating audio prompts such as "the keys are on the hook in the hallway," and/or speech-to-text options as well as accommodations for those who are hard-of-hearing such as blinking lights or large motions when the tablet's audio will not suffice. Finally, in order to facilitate adoption of this robotic assistance into an older adults' everyday routines, the robot would ideally be able to integrate with other technologies such as Google Home (www.store.google.com/us/product/google_home) or Alexa (<https://www.amazon.com/Amazon-Echo-And-Alexa-Devices>).

References

Alexa (<https://www.amazon.com/Amazon-Echo-And-Alexa-Devices>)

Alzheimer's Caregiving: Changes in Communication Skills, n.d (national institutie on aging)

Dermody, N., Hornberger, M., Piguet, O., Hodges, J. R., & Irish, M. (2015). Prospective memory impairments in Alzheimer's disease and behavioral variant frontotemporal dementia: clinical and neural correlates. *Journal of Alzheimer's Disease*, 50(2), 425–441. doi:10.3233/jad-150871.

Estimotes; Contact Estimote, Inc., San Francisco, CA

Feingold-Polak, R., Elishay, A., Shahar, Y., Stein, M., Edan, Y., & Levy-Tzedek, S. (2018). Differences between young and old users when interacting with a humanoid robot: a qualitative usability study. *Paladyn, Journal of Behavioral Robotics*, 9(1), 183–192. doi:10.1515/pjbr-2018-0013.

Goher, K.M., Mansouri, N., & Fadlallah, S.O. (2017). Assessment of personal care and medical robots from older adults' perspective. *Robotics and Biomimetics*, 4(1). doi: 10.1186/s40638-017-0061-7.

Google Home (www.store.google.com/us/product/google_home)

Hernandez Cardenache, R., Burguera, L., Acevedo, A., Curiel, R., & Loewenstein, D.A. (2014). Evaluating different aspects of prospective memory in amnesic and nonamnesic mild cognitive impairment. *ISRN Neurology*, 2014, 1-7. doi: 10.1155/2014/805929.

Hone, K.S., & Graham, R. (2000). Towards a tool for the Subjective Assessment of Speech System Interfaces (SASSI). *Natural Language Engineering*, 6(3-4), 287-303. doi: 10.1017/S1351324900002497.

- Hsu, Y., Huang, C., Tu, M., & Hua, M. (2015). Prospective memory in subjective cognitive decline: A preliminary study on the role of early cognitive marker in dementia. *Alzheimer Disease & Associated Disorders*, 29(3), 229-235. doi: 10.1097/WAD.0000000000000060
- Jiang, J., Huang, Z., Huo, B., Zhang, Y., & Song, S. (2018). Research progress and prospect of nursing robot. *Recent Patents on Mechanical Engineering*, 11(1), 41-57. doi: 10.2174/2212797611666180306124236
- Kashi, S., & Levy-Tzedek, S. (2018). Smooth leader or sharp follower? Playing the mirror game with a robot. *Restorative Neurology and Neuroscience*, 36(2), 147–159. doi:10.3233/rnn-170756.
- Kittmann, R., Fröhlich, T., Schäfer, J., Reiser, U., Weißhardt, F., & Haug, A. (2015). Let me introduce myself: I am Care-O-bot 4, a gentleman robot. In: Diefenbach, S., Henze, N. & Pielot, M. (Hrsg.), *Mensch und Computer 2015 – Proceedings*. Berlin: De Gruyter Oldenbourg. (S. 223-232).
- Martins, M., Santos, C., Frizera, A., & Ceres, R. (2015). A review of the functionalities of smart walkers. *Medical Engineering and Physics*, 37(10), 917-928. doi: 10.1016/j.medengphy.2015.07.006.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. doi:10.1111/j.1532-5415.2005.53221.x.
- Pino, M., Boulay, M., Jouen, F., & Rigaud, A. (2015). “Are we ready for robots that care for us?” Attitudes and opinions of older adults toward socially assistive robots. *Frontiers in Aging Neuroscience*, 23(7), doi: 10.3389/fnagi.2015.00141.
- Pollack, M. E., Brown, L., Colbry, D., Orosz, C., Peintner, B., Ramakrishnan, S., . . . McCarthy, C. E. (2002, Jul 29). Pearl: A Mobile Robotic Assistant for the Elderly. Paper presented at the Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care. AAAI 2002, Edmonton, CA.
- Qualtrics (2005). Version April, 2018. Provo, Utah, USA. Retrieved from <https://www.qualtrics.com>.
- Smarr, C., Mitzner, T.L., Beer, J.M., Prakash, A., Chen, T.L., Kemp, C.C., & Rogers, W.A. (2015). Domestic robots for older adults: Attitudes, preferences, and potential. *International Journal of Social Robotics*, 6(2), 229-247. doi: 10.1007/s12369-013-0220-0.
- Stafford, R.Q., MacDonald, B.A., Li, X., & Broadbent, E. (2014). Older people’s prior robot attitudes influence evaluations of a conversational robot. *International Journal of Social Robotics*, 6(2), 281-297. doi: 10.1007/s12369-013-0224-9.

- Tam, J.W., Van Son, C., Dyck, D., & Schmitter-Edgecombe, M. An educational video program to increase aging services technology awareness among older adults. *Patient Education and Counseling*, 100(8), 1564-1571. doi: 10.1016/j.pec.2017.03.020.
- Wang, R.H., Sudhama, A., Begum, M., Huq, R., & Mihailidis, A. (2017). Robots to assist daily activities: Views of older adults with Alzheimer's disease and their caregivers. *International Psychogeriatrics*, 29(1), 67-79. doi: 10.1017/S1041610216001435.
- Wilson, G., Pereyda, C., Raghunath, N., de la Cruz, G., Goel, S., Nesaei, S., ... Cook, D.J. (2019). Robot-enabled support of daily activities in smart home environments. *Cognitive Systems Research*, 54, 258-272. doi: 10.1016/j.cogsys.2018.10.032.
- Wu, Y.W., Cristancho-Lacroix, V., Fassert, C., Faucounau, V., de Rotrou, J., & Rigaud, A. (2016). The attitudes and perceptions of older adults with mild cognitive impairment toward an assistive robot. *Journal of Applied Gerontology*, 35(1), 3-17. doi: 10.1177/0733464813515092
- Yu, R., Hui, E., Lee, J., Poon, D., Ng, A., Sit, K., ... Woo, J. (2015). Use of a therapeutic, socially assistive pet robot (PARO) in improving mood and stimulating social interaction and communication for people with dementia: Study protocol for a randomized controlled trial. *Journal of Medical Internet Research*, 4(2). doi: 10.2196/resprot.4189.