International Journal of Social Robotics (2018) 10:621–641
https://doi.org/10.1007/s12369-017-0456-1

Supporting Human Autonomy in a Robot-Assisted Medication Sorting Task

Jason R. Wilson[1,2] · Nah Young Lee[1] · Annie Saechao[1] · Linda Tickle-Degnen[1] · Matthias Scheutz[1]

Accepted: 7 December 2017 / Published online: 19 February 2018

Abstract
Medication management is a significant challenge for older adults, and the resultant drug-related problems are linked with hospitalizations and increased need for nursing homes. In this work, we explored the role of a socially assistive robot for one aspect of medication management: sorting. Specifically, we proposed a human-centric approach towards the design of a robot assisting in a medication sorting task. The approach is based on the analyses of occupational therapists who are trained in evaluating and assisting older adults in important self-care skills and emphasizes the role of autonomy on the part of the person performing a medication sorting task. We developed and evaluated two robot prototypes that assist a person in a medication sorting task. In both prototypes, evaluated by students (N = 31) at an American university, we found that subjects voluntarily greeting the robot experienced the emotion of the interaction differently from non-greeters. Greeters of the physical robot gave a lower emotional rating of the interaction, whereas greeters of the virtual robot found the emotion of the experience to be better than the non-greeters.

Keywords Socially assistive robot · Medication sorting · Medication adherence · Autonomy · Occupational therapy

1 Introduction

Many older adults need to take multiple medications each day, and it can be challenging to consistently and accurately follow all of the prescriptions. Following the instructions that are given for prescribed medications is referred to as *medication adherence* [20,36], and failure to follow the instructions is known as *medication non-adherence*. Unfortunately, medication adherence rates tend to be low across many chronic conditions with complex medication schedules [14] (e.g., Parkinson's disease [32], dementia [33], hypertension [9]), and the consequences of non-adherence range from hospitalization [45] to needing to live in a nursing home [28]. A 1995 report said that in 1 year, there was over $76 billion in costs for drug-related problems in the United States, with almost 9 million hospitalizations costing over $47 billion and over three million older adults were placed into long-term care facilities with $14 billion in costs [24]. A 2001 update on this report revealed that the problem has not improved, with over $177 billion in costs for drug-related problems, of which $121.5 billion was related to hospitalization and $32.8 billion was for admissions to long-term care facilities [15]. Non-compliance is not a problem unique to the United States, as another 2001 report found that 26% of admissions at an Australian hospital were the result of non-adherence [11]. We can avoid some of these costs and help older adults live in their own homes longer by increasing their ability to manage their medications.

Strategies for improving medication adherence include using cues to remind people to take their medications and using pill boxes to organize daily doses [36]. While pill boxes are commonly suggested to improve medication adherence (e.g., [3,31]), sorting medications into a pill box introduces new challenges because it requires cognitive abilities such as verbal memory, cognitive flexibility, and executive functioning [46], and even mild challenges in these areas lead to non-adherence in older adults [10,14,19]. To help in the sorting process and thus improve medication adherence rates, one may get assistance from a spouse, a friend, a nurse, or some other person. These care partners provide critical assistance, but it creates a demand for people to assist. With the

This project was in part supported by NSF Grant #IIS-1316809.

✉ Jason R. Wilson
wilson@cs.tufts.edu

1 Tufts University, Medford, MA, USA

2 Northwestern University, Evanston, IL, USA

growing number of older adults, this demand for people to assist is growing, and we need to consider how various forms of technology could help fill the gap between the number of people that need assistance and the number of people that can provide assistance. Some have argued that technological aids can fill this critical role and help meet the growing demand [8,31,37,38]. We need to continue to investigate how robots and other technologies can be used to supplement the care and assistance provided by humans, and particularly we must consider how technologies can address issues related to medication management.

An automated reminder systems is an example of a technology that is being used to improve medication adherence. Medisafe is one such automated reminder system that provides alerts on a person's smartphone when it is time to take a pill and indicates which medication is to be taken, and Medisafe has been shown to improve adherence rates [52]. Many other smartphone apps have been evaluated and recommended for improving medication adherence [13]. Others have designed robots to provide medication reminders (e.g., [12,54]), as some people have indicated a preference for robots over humans to provide medication reminders [43]. As for helping in taking the correct medications, there are many devices that will unlock a compartment containing the pills to be taken at a particular time (e.g. [6,40]). While all other compartments remain locked, the person is only able to access the pills that are to be taken at that time. These technologies are an evolution of the basic pill box, but they still have some of the same challenges in requiring a person to sort the pills into the device. In many cases, these devices also then need to be programmed, thus leading to an increasing burden on the person providing assistance. We suggest that instead of requiring more assistance from another person, we need to enable the person taking the medications to be able to provide more care for him or herself. We take the idea of having robots assist in medication reminding and applied it to another aspect of medication management, medication sorting. We expect that a robot would have advantages over a smartphone, as the robot's physical embodiment and presence have been shown to have an effect in other applications [30,48,53].

We have designed a robot that provides cognitive and social support to a person sorting medications. Instead of providing direct physical support to the person, the robot provided only enough assistance to allow the person to successfully complete the medication sorting on their own. By providing minimal amount of assistance, the person may retrain control and responsibility for an important health management activity. To accomplish this, we used a structured framework of increasing assistance [39], so that the robot provided the amount of assistance that met the need of the person. By enabling the person to take the necessary actions to successfully sort the medications, we allowed the person to feel in control and contributed to the person's sense of *autonomy*. In this paper, we describe further the concepts of *autonomy* and *dignity* from the perspective of occupational therapists who are regularly tasked with evaluating and assisting older adults with self-care skills. Next, we present the medication sorting task and discuss how a social robot may assist with it. Finally, we describe the designs of two prototypes and discuss user evaluations of the system.

2 Background

There are two main areas of socially assistive robots [16] in a healthcare/eldercare environment: *companionship* and *service* [8]. Examples of robots that work on social commitment and companionship are Paro [49,51] and AIBO [5,25,47]. They were designed to use multifaceted sensory stimulations (e.g., auditory, tactile, and visual) to help interact and engage with elders, increase mood and vigor [51], reduce stress [49], and decrease loneliness [25]. Moreover, there is preliminary evidence that these robots also improved staffs' and caregivers' overall well-being by decreasing the feeling of burnout [51].

Service robots have features to support navigation, mobility, and safety, and remind users of routine activities (e.g., toileting, eating, cooking, taking medication) [8]. They are able to provide cognitive aids to elders to monitor performance as well as individualizing to the elders' preferences, constraints, and time of performing the activities. An example of a service robot is Pearl, which reminded people about routine activities and provided navigation guidance through their environments. Another example of a robot assisting in navigation used an autonomous robot to assist a walking group of older adults with dementia [21]. Other examples of service robots are iCat [22] and Care-O-bot [18], or general purpose robots like the PR2 [38] or the Pioneer [34] used as service robots. Each of these robots use social interaction as the primary mode of assisting older adults, and similarly our work would be considered a service robot that provides social assistance. We also seek to gain some of the benefits associated with companion robots (e.g., improved mood, decreased loneliness) by paying attention to the emotional aspects of the interaction. The robot should contribute to a pleasant environment and be emotionally supportive.

3 Principles of Occupational Therapy

In designing a robot to assist older adults, few have incorporated the perspective of occupational therapists who are trained in assisting older adults in their daily activities. Occupational therapists (OTs) enable individuals to participate in everyday life activities (occupations), whether it be physical, mental, social, sexual, political, or spiritual [55]. The *Performance Assessment of Self-care Skills* (PASS) manual defines 26 activities of daily living (ADL) [39]. Some ADLs, such as toileting or dressing, are fundamental aspects of self-care. Other ADLs may require more cognitive capabilities and are referred to as instrumental activities of daily living (IADL). These activities include cooking, shopping, and managing medications. As part of the process of engaging the individual in these activities, the actions of an OT go beyond supporting the physical demands of any activity. Throughout the process, an OT must remain cognizant of the physical, social, personal, and cultural contexts of the activities [1].

A major emphasis of occupational therapy practice is a client-centered approach of assessing the client holistically (physically, emotionally, socially, psychologically). Each client brings his or her own unique background and perspective, and occupational therapy is a collaborative process that is client-driven to meet the needs of each individual [41]. Among the core values of occupational therapy is maintaining the dignity of the individual by treating the person with respect during all interactions. Similarly, one of the principles and standards of conduct for OTs is *autonomy*—respecting the individual's choice and confidentiality. The autonomy of the individual includes having minimal dependencies on people, devices, or technologies and being able to freely make choices. The individual has a right to make decisions based on the direct care of their own health [2].

The importance of autonomy and the right to make decisions applies to all the activities of daily living, including medication management. As much as possible and as long as an individual demonstrates the necessary cognitive, social, and emotional capabilities, an individual should be free to exercise his or her own judgment in how to adhere to a medication regime, what is the best time for a medication that fits the individual's schedule, when to see a doctor, what to have for dinner, and other decisions relevant to activities of daily living. An OT should respect the choices made by an individual while still providing assistance to guide the person to follow essential medical plans and improving health and quality of life.

One way to maximize the autonomy of the individual is to minimize the amount of assistance provided (see Fig. 4). As assistance increases, the individual may grow dependent on the assistance which can reduce how much the individual is self-reliant, i.e., how much the individual uses his or her own capabilities (physical, cognitive, or otherwise) to execute a task. This dependency restricts or constrains the individual, thus limiting his or her autonomy. Conversely, minimizing the assistance requires a person to rely on his or her own capabilities and allows the individual to freely decide and act. However, we note that too little assistance can also have a negative effect on autonomy. Thus, is it important to match the level of assistance to the capabilities of the individual to maximize autonomy.

In assessing the ability of a person to perform self-care activities, OTs use a scale of types of assistance necessary to complete an activity. The PASS manual defines 9 levels of assistance, ranging from verbal support at level 1 to complete assistance at level 9. "As the assists given progress from Level 1 to Level 9, the patient assumes less responsibility and the therapist assumes more responsibility for independent, safe, precise, or client task performance. Therefore assists are given only if they are needed to progress task performance or to ensure safe performance" [39]. The PASS manual also provides a framework for assessing an individual's ability to perform self-care skills, like sorting medications and bathing. Any activity that requires no support, not even a level 1, indicates that the person is fully capable of self-care for that activity. If the individual occasionally needs some level 1–6 assistance, then self-care may be acceptable but improvements can be made [39].

As already mentioned, being able to perform self-care skills requires a certain amount of autonomy on the part of the individual, but providing assistance impacts the person's autonomy and dignity. Therefore, in order to maximize autonomy, we plan to match the level of assistance to the capacities of the individual, to not over- or under-assist, rather to match assistance to what is needed.

4 Medication Sorting Task Analysis

One part of medication management is sorting medications, which entails organizing pills according to the day of the week and time each pill is to be taken. Each pill is placed into a pill box or on a sorting grid, as shown in Fig. 1. The grid contains columns for each day of the week and rows for different times of day. Four subdivisions for time of day (Morning, Noon, Evening, and Bedtime) can be used.

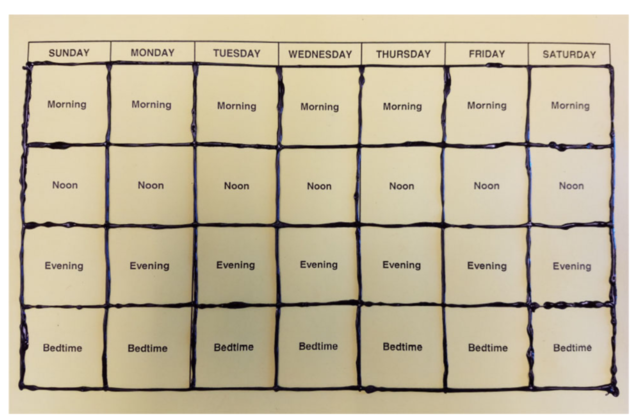

Fig. 1 The medication grid has columns for each day and four rows for morning, noon, evening, and bedtime

To design a robot to assist with medication sorting, we first needed to better understand the steps involved in the process of sorting medications. For this, we turned to a standardized protocol for organizing pills onto a medication grid according to the instructions on the pill bottles [39]. In this protocol, the therapist orients the client to the task, handing the person a bottle and asking them to read the label. Then the person distributes the pills in the medication grid for the current day and the next day. The same steps are then taken to sort the pills of a second medication. The protocol is intended to be flexible and modifiable according to the needs of researchers, health practitioners and health-care clients/patients. For example, the protocol may still be followed if the person does not have two medications by substituting vitamins in place of a medication.

In taking a more human-centric approach to designing the robot, it was important to understand not just the physical dynamics of a task but also the emotional, social, environmental, and safety concerns relevant to the task. To this end, we developed an in-depth task analysis of the medication sorting activity. The task analysis broke down the task into motor, social, and cognitive domains of performance.

Based on the outlined procedure in the PASS manual, we also produced a short video simulation of a medication sorting activity. One student played the role of the patient sorting the medications, and another student provided the instructions and assistance of the task. This baseline scenario would later be used to create the outline of the script used to program the robot.

4.1 Evaluation of Task Analysis

To evaluate the task analysis and video simulation, we gathered four trained occupational therapists to review the material and provide feedback. We prepared a 17 item Likert-scaled questionnaire asking about the realism, comprehensiveness, and relevance of the content of the task analysis and the video simulation. Each of the experts reviewed the task analysis, watched the video, and then independently completed the questionnaire.

After everyone had completed the questionnaire, author LTD led the group in a discussion regarding the medication sorting task, the task analysis, and how a robot could assist with medication management. To facilitate discussion about the robot and its capabilities and possible roles, we presented pictures of the Nao robot as an example of a small robot that could be used for social assistance in a task like medication sorting. The focus group discussion was audio recorded, and the recordings were later transcribed.

During the discussion, the team of experts agreed that

- the task analysis was comprehensive and complete
- the video simulation lacked realism and complexity, and
- some of the physical details of the robot are important to consider.

Comprehensive Task Analysis The team of experts agreed that the medication sorting task was a valuable model of daily life medication management and that the content of the activity analysis was comprehensive and complete. The main criticism was that the safety concerns were not sufficiently represented.

Realism and Complexity of Task The experts found the controlled and simplistic lab setting of the video simulation unrealistic and inconsistent with the home environment, where the task would typically be done. The experts suggested there should be a tray in which to pour the pills, which could help prevent the pills from getting lost in the clutter of the table or from rolling off the table. Another inconsistency was that the video demonstrated the medication grid positioned between the two people, but it would typically be positioned in front of the person doing the sorting.

In addition, the OT experts commented that there were areas of the task that lacked complexity. The simulation used only two pills and only 2 days of the week. While this is exactly what is described in the PASS, it was found to be not representative of common practice. In the prototypes of the robots that are described in the rest of this paper, pills needed to be sorted into all 7 days of the week.

Physical Details of Robot The OT experts made several suggestions regarding the robot's size, position, and gesture capabilities. A robot that is too big or has an eye position above the person could be intimidating or viewed as having power. The Nao robot that we presented was well-liked because it was small and can be kept at eye level. Another advantage of the Nao robot is it can make hand gestures. Some thought that gestures like pointing at a lost or misplaced pill would be helpful for a person who may not be able to see or hear well.

5 Initial Robot Prototype

Based on the detailed task analysis and the feedback from the OT experts, we developed an initial prototype of a robot for assisting in a medication sorting task. Motivated by the need to support a person's autonomy, as defined as an important OT principle, we developed a plan to have the robot provide assistance that matched the need of the person. This section describes the design choices made, the evaluation of the robot in a setup that is more complex and realistic than in the video simulation, and discusses the results of the evaluation and design considerations to be further explored.

The primary questions in designing this prototype were as follows: (1) how should the robot decide what level of assistance to provide? (2) for each level of assistance, what does it mean to give that type of assistance? and (3) what are the relevant dimensions for evaluating the robot and people's experience with the robot?

To address these questions, we defined the following goals in designing this robot prototype:

– **Follow Established Protocol** The task analysis and video simulation provided the necessary content for how a medication sorting task is done. We did not want to introduce any changes to the procedure and looked to design the robot to assist in such a task just as is done in actual practice.
– **Provide Cognitive and Social Support** The robot should provide cognitive and social support but not direct physical support. The robot must not touch the pills or the person at any time. See Sect. 5.1.3 for why the robot was limited to cognitive and social support.
– **Support the Autonomy of the Person** The assistance the robot provides must support the autonomy of the individual by matching the need of the person with the assistance provided.
– **Not Requiring Any Human Operators** The robot must not require any human operators and must operate fully autonomously. We are ultimately interested in building a robot that can be used in a practical application, and using remote operators in those application domains would defeat the purpose.

5.1 Architecture

To accomplish these design goals, the necessary components of the robot architecture were vision, action script execution, medication management and assistance, and robot control (see Fig. 2). These components were implemented using ADE, which is the middleware for the DIARC robot architecture [42].

5.1.1 Vision

To be able to observe the human while sorting medications, the robot needed to be able to monitor the medication grid to determine the progress of the task. The vision component was responsible for perceiving the environment so that other components may use this information to infer the state of the task and of the user. The Vision component used the cameras on the robot to capture the state of the medication sorting grid (shown in Fig. 1) and the sorting tray, reporting how many pills of each type are in each cell in the grid. The type of pill was determined by the pill color. The sorting tray, which was a small circular tray, was treated as a separate grid with one cell. The vision system could also report how many pills of each type were in the tray.

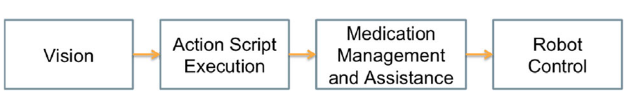

Fig. 2 The flow of control amongst the components used to implement the medication sorting robot

During the process of identifying the number and position of pills on the grid and tray, the Vision component also determined if the number of pills has changed. A change in the number of pills in any of the cells of the grid or tray was used by the other components to infer whether an action might have been taken by the person.

5.1.2 Action Script Execution

The action script execution component was responsible for managing the sequence of actions the robot took, and thus enabled the robot to act without the intervention of any remote operators. The robot's role in a medication sorting task was encoded as an *action script*, or a sequence of actions that the robot executed. Each action in the script may be another action script, thus creating a hierarchy of scripts. For example, we constructed a top-level script that was responsible for managing the overall flow of the robot's actions:

1. Initialize.
2. Provide the task instructions.
3. Assist in sorting the pills of medication 1.
4. Assist in sorting the pills of medication 2.
5. Finalize.

Each of these steps was another action script. The first and last actions were mostly bookkeeping actions designed to manage the information about the pills and to start and shutdown the Vision component. As part of the first action, the medication management and assistance component was informed of the goal state of the medication sorting task. The goal state consisted of the number and position of each of the pills in the medication grid. The action script for step 2 included the robot introducing itself and orienting the person to the task by pointing out the grid and the containers of pills.

Steps 3 and 4 refer to the same action script that defined the actions the robot took to assist in the sorting of one medication. Each reference to this action script included an identifier for the sorted medication. The steps of this action script are shown in Fig. 3. The action script started by having the robot say which pill is to be sorted next, instruct the person to pour some pills into the tray, wait for pills to appear in the tray, and then provide the instructions for that medication. The next part of the action script was the main loop, in which the

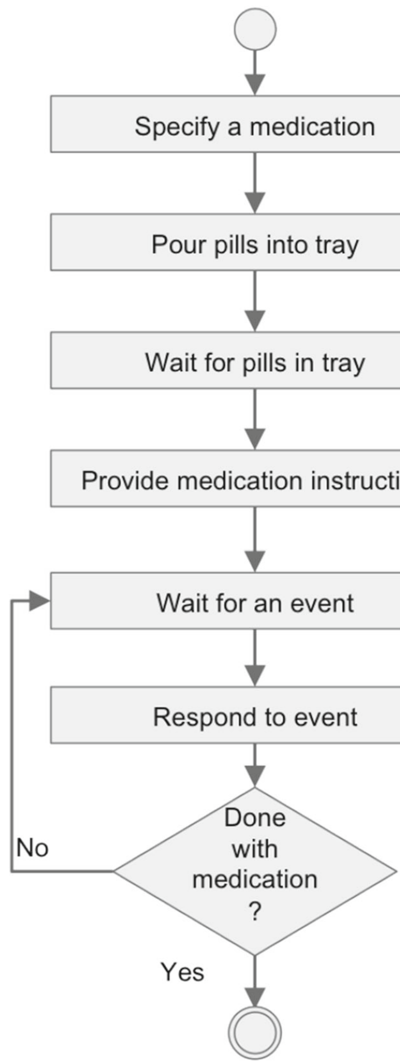

Fig. 3 The flow of the action script for assisting with the sorting of the pills for a given medication. The action script starts with the robot specifying which medication is to be sorted next. When all the pills for that medication have been sorted, the action script ends and control is returned to the calling action script

actions for the robot were for it to wait for changes in the state of the grid, respond to the state change, and check if all the pills for this medication have been sorted. The waiting for a state change also had a timeout, which allowed the robot to respond to the event where no change in the state of the grid was detected, but the robot still needed to provide the user with some assistance. Responding to the event was handled by the medication management and assistance component.

5.1.3 Medication Management and Assistance

The medication management and assistance component was responsible for providing the appropriate assistance given the state of the task (e.g., Fig. 4). This was done in a two-step process. First, it generated an event representing the state of

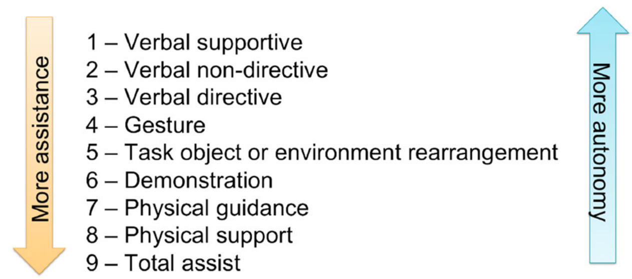

Fig. 4 The 9 levels of assistance as defined in the PASS manual [39]. A problem with providing too much assistance is that autonomy may decrease

the task and then selected the appropriate action based on the state of the task. Once the action had been selected, it sent the information about the action to the robot control component for execution (see Sect. 5.1.4).

Event Generation The first step in determining the appropriate assistance for the robot to give was to construct an event representing the state of the task. Based on the observations provided by the Vision component, the event included the number and position of the pills in the medication sorting grid and tray. The event also included information about how the current state of the medication grid was related to the specified goal state, how the current state of the medication grid was different from that specified in the previous event, whether this event represented any delays or hesitances the person may have been having, and whether any errors had been made recently. Each of these are briefly explained here.

To recognize whether the person was on course or whether the current state of the medication grid had any mistakes, the component compared the observed state produced by the Vision component with the desired goal state. This required the component to have the knowledge of the goal state, which had the correct allocation of pills to the cells in the medication sorting grid. This knowledge was provided to the medication management and assistance component when the *Initialize* action was executed by the action script execution component. If any cell of the grid in the current state had more pills of a particular type than was specified in the goal state, then the current state was inconsistent with the goal. In the case of inconsistencies, the location of the misplaced pill was logged in the event.

To identify whether an action had occurred, the component compared the most recent observed state with the previous one. If any pill on the grid or the tray had been added, moved, or removed, then an action had occurred. If there was no change in the grid or tray, then the event was the result of a timeout and tagged as representing a hesitation.

If an action had occurred, then it needed to determine if the action was a correct one or not. An action was considered correct if (1) a pill had been *added* to a position in the grid, and the goal state indicated there should be a pill there, and/or

(2) a pill had been *removed* from a position in the grid and the goal state indicated there should not be a pill there.[1]

The final step was determining if there had been numerous recent errors made. To determine this, the component scanned over the recent events to find how many of them represented some form of error. An error was either an incorrect action or a hesitation. For this prototype, an event was considered recent if it was one of the five most recent events. Thus, the number of recent errors was the number of events out of the past five that represented an error.

Assistance Selection In accordance with our goal to support the autonomy of the person doing the medication sorting, the robot must provide assistance that match the need of the person. To accomplish this, we structured the decision process around the levels of assistance as defined in the PASS [39]. In focusing on cognitive and social support rather than physical support, we limited the robot to only being able to provide the first 4 levels of assistance. Starting at level 5, the assistance may require physical contact with the person being assisted. There are ethical concerns when it comes to personal touch between a robot and a person. Limiting the robot to the first four levels emphasized the cognitive and social interaction while minimizing the risk of any ethical or moral violations.

We organized the decisions the robot makes in selecting the appropriate action into a decision tree. We engineered the tree such that the depth of the tree represented the extent to which the person experienced challenges and needed more assistance is necessary. Figure 5 shows the high-level decisions and the corresponding level of assistance. The remainder of this section describes each level of assistance.

Level 0. We started with a baseline of a level 0 assistance, which was selected if the current state of the medication grid was consistent with the goal state. The PASS does not define a level 0, but we found it necessary for the robot to provide some basic acknowledgment of what the person is doing and allow the person to know that the robot is watching and correctly functioning. Having the robot provide simple feedback was consistent with a natural interaction a human would have in assisting with the task. In watching the video simulation we produced, we noticed the person playing caregiver role regularly provided back-channeling signals. This was done even though the person was trying to give no assistance. Also, we have found that if the robot had no reaction to what the person was doing, it was hard to tell if the robot was working. As a result, we introduced this level 0 assistance to provide basic feedback, in which the robot may give a short verbal utterance (e.g., "Good") or more subtle responses like a head nod or a blink of the eyes.

[1] We recognize that this definition of a correct action is not complete, but it was sufficient for this prototype and was improved in the next prototype.

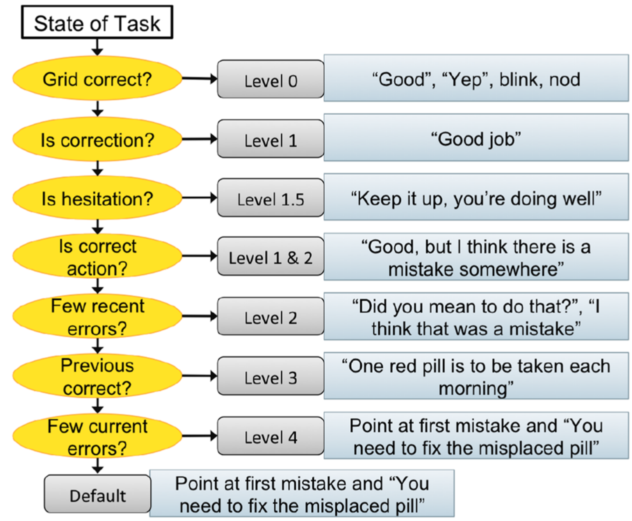

Fig. 5 Decisions for determining level of assistance and examples of each level. The yellow ovals are queries about what the robot has observed. If the result of the query is *true* then the response on the right is generated. Otherwise, the next decision down is tested

Level 1. A level 1 assistance was intended to be *verbal supportive*, providing encouragement for the person to continue or to complete the task [39]. There were three conditions under which the robot would select a level 1 assistance: (1) If the action that the person has taken was a correction of a misplaced pill, then a level 1 response was generated (e.g., the robot says "Good job"); (2) If the most recent event was a hesitation event, then the selected action was one that we had categorized as a level 1.5. An example of a level 1.5 assistance was for the robot to say, "Keep it up, you're doing well." This statement was intended to encourage the person to take another action; (3) If the most recent event represented a correct placement of the pill, but there were other pills on the medication grid that were incorrect, then the robot responded with a statement that was categorized as both level 1 and 2 (e.g., "Good, but I think there is a mistake somewhere.")

Level 2. The PASS defines a level 2 assistance as *verbal non-directive* [39], and we used this to provide some verbal indication that a mistake had been made without directly referencing what the mistake was or how to fix it. A level 2 assistance was selected when there had been few recent errors. For this prototype, we defined "few" recent errors as one error in the past five events. Zero errors were also constituted as "few," but at this point in the reasoning, we would have already known that the most recent action was incorrect, and thus there must have been at least one error. When there were few recent errors and a level 2 assistance was selected, the robot responded with an utterances like "Did you mean to do that?" or "Try a different way."

Level 3. Level 3 was *verbal directive* assistance, which provided more direct information about how to complete the task. A level 3 assistance was selected when the above conditions did not apply and the previous state of the medication grid was consistent with the specified goal state. For this prototype, when providing a level 3 assistance, the robot repeated the instruction for the medication currently being sorted.

Level 4. A level 4 assistance included *gesture*, a non-verbal communication to provide more information to the person. If there were several errors in the current grid, then a level 4 assistance was selected. For medication sorting, the appropriate gesture was for the robot to point at the cell in the grid in which there was a mistake. This was done along with the robot saying, "You need to fix the misplaced pill."

Finally, if none of the above conditions applied, a default response was given since the person needed more assistance, but the robot could not give more than a level 4 assistance. In this case, the robot responded with, "You seem to be having great difficulty. Take your time and think carefully."

5.1.4 Robot Control

The robot control component provided the interface between the DIARC architecture and the low-level software controlling the Nao robot. Included in this interface were the mechanisms for making the robot produce speech from text, move its arm so that it could point, nod its head, and blink its eyes. The interface directly provided each of these capabilities. For example, if a level 1 assistance was selected, the procedure call `say("Good")` caused the robot to produce the corresponding speech output.

5.2 Evaluation

In evaluating this initial prototype we wanted to evaluate how people perceived the robot as it provided assistance. We were interested in the impressions people had of how well the robot functioned, the experience interacting and being assisted by the robot, and whether it was supportive and pleasant. The evaluation was also an opportunity to test the quality of the fully autonomous robot in a constrained but unscripted task.

To evaluate the prototype, the robot assisted a person in completing a medication sorting task. In order for the task to have more ecological validity for human application, the task needed to be slightly more complex than featured in the video simulation. We chose to continue to require the person to organize only two medications, but the pills needed to be allocated to all 7 days of the week. Unlike the video simulation, we did not want the medication sorting to be completed without the need for assistance. Instead, the robot needed to be able to provide assistance both when a pill is misplaced and when the person is slow to complete the next step in the task.

5.2.1 Evaluation Setup

Participants Students (N = 11) from Tufts University were recruited by word of mouth to participate in a human–robot interaction study. Most were affiliated with the occupational therapy program, but no demographics were collected from the participants. There were three participants who were involved in developing the analysis of the medication sorting task but did not know any details about the design and implementation of the robot prototype. We chose to evaluate our initial designs with students to allow us to rapidly prototype and evaluate different designs.

Methods and Design Each participant completed an informed consent statement and then interacted with the robot to complete the medication sorting task and then completed a 20-item questionnaire. We recorded a video of each participant completing the task with the robot.

The setup used a Nao robot positioned on the table across from the participant, a medication grid (as seen in Fig. 1) and tray, and two cups of simulated medicines (candies). A research assistant informed each participant that the robot will be assisting in a task involving placing the medications onto the medication grid. Each participant was also instructed to follow the instructions of the robot while intentionally making some mistakes. Two types of mistakes were requested of each participant. First, the person was to hesitate in acting, taking more than 3 s to place the next pill. The 3 s threshold was easily adjustable and would probably need to be a longer duration in future prototypes, but we found in pilot testing that 3 s was a sufficient amount of time for the purposes of our evaluation. The other type of mistake was to incorrectly place a pill. Any mistake may be made more than once if the participant chose to do so.

We gave the participant these guidelines for making mistakes in an effort to encourage a more natural interaction with the robot. An alternative would have been to give the participant a script to follow, where the script precisely indicated what the person was to do, including when and where to make mistakes. The concern was that the participant might be more focused on following the script than on the interaction with the robot. Another alternative was to give the participant no guidelines. Since the task was very simple, it would be unlikely that the participants would make any mistakes. This would not have allowed us to test the feedback and assistance that the robot gave. Providing the participant guidelines in which to operate was a reasonable compromise that allowed the participant to naturally interact with the robot while still making mistakes so that we could evaluate the robot's performance.

In the questionnaire that followed the medication sorting, the participants completed 19 Likert-scaled questions on a 5-point scale ranging from "strongly disagree" to "strongly agree" and one open question for general comments. The subject of the questions were intended to address whether people found the robot to efficiently perform its function, be trustworthy, be supportive, build rapport, promote safety, respectful of social context, and support a positive experience. Trust and reliability questions were based on [35]. Questions regarding the robot being supportive were adapted from [23]. Rapport questions came from [7]. The questions about the mood of the experience were derived from the affect dimensions described in [44]. In preparation for future studies with the target population, we also collected information about how the robot may be perceived by one's family or care providers and other factors contributing to the desirable experience (e.g., the person feeling in control and responsible). All of the questions are in the "Appendix" section.

5.2.2 Analysis

The questionnaire item responses were analyzed with descriptive statistics (see the "Appendix" section for the full questions). Before any analysis, we inverted the answer values for question 10 due to it being negatively framed. In addition to the descriptive statistics, we also compared the set of answer values for each question to the mean of all the question answers using a one-sample t test. The mean of all question answers ($M = 3.837$) is used as the hypothesized mean in the one-sample t test to adjust for the positive bias in the questionnaire answers. Lastly, we conducted a principal component analysis (PCA) to find the dimensions that best described the questionnaire results. We tried the PCA using 3, 4, 5, 6, and 7 components.

5.2.3 Evaluation Results

The means of each of the questions are shown in Fig. 6. The results indicated that people felt in control during the task, felt that they understood what was happening, and felt responsible for completing the task. This was demonstrated in the high scores for 17–19. As intended, the low scores for question 7 reflect that they also did not feel physically supported during the task. The low score for question 10 was likely due to the confusing wording of the question.

Some participants had previous experience with the project, having worked on the task analysis and video simulation, but had no experience with the robot. We compared the results of those participants with the rest to see if there were any differences. Only questions 7 and 8 (both relating to how much the robot supports the person) had a statistical difference between the groups: 7 ($t(7) = 3.33$, $p = 0.01$) and 8 ($t(7) = -2.57$, $p = 0.04$). For both questions, participants with previous experience with the project reported less emotional and physical support compared to participants who did not have this experience.

In assessing the components produced from the PCA, the analysis producing 3 components was not only the simplest but perhaps also made the most sense conceptually. The first component described the robot in terms of its proper functioning, safety, and trustworthiness. We label this component *effectiveness*. The second component described the experience with robot as inclusive, helpful, and clear, and we refer to it as *supportive assistance*. The third component described the support the robot provides for the emotion and mood of the person. This component is simply labeled *emotion*. How

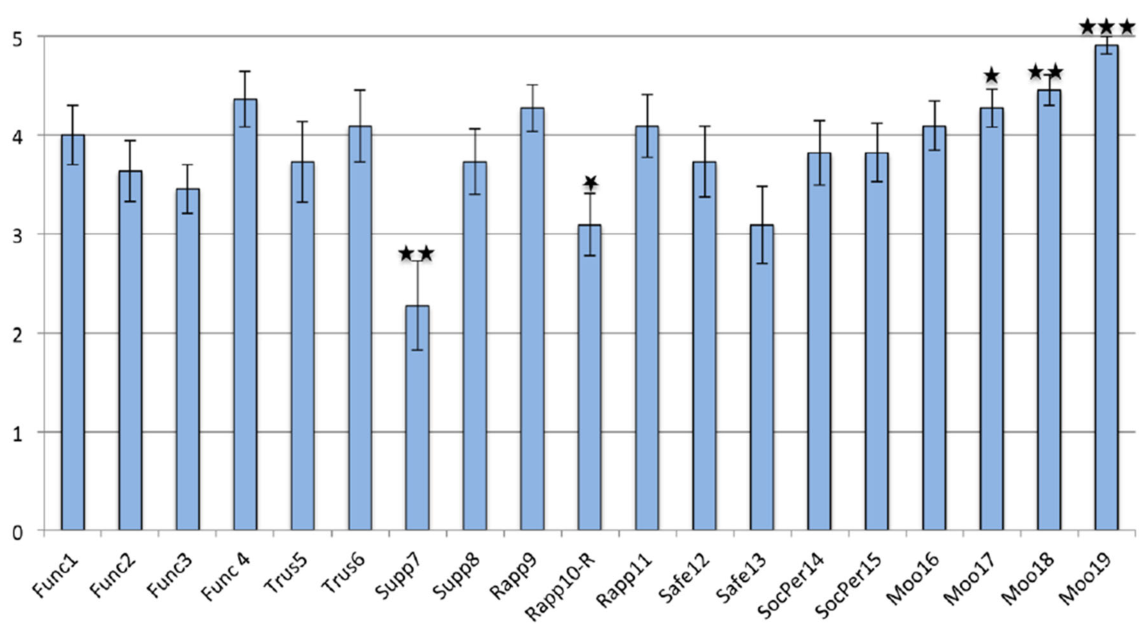

Fig. 6 The means and standard error of each of the questions. Some questions significantly differed from the mean of all the questions ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$)

Table 1 Loading of questions onto each component

Question	Effectiveness	Supportive assistance	Emotion
Trus6	0.924		
Safe12	0.812		
Func2	0.809		
Safe13	0.777		
Trus5	0.769		
Rapp9	0.695		
SocPer14	0.645		
SocPer15	0.602		
Func3	0.571		
Func1		0.845	
Mood18		0.803	
Rapp11		0.646	
Mood19		0.539	
Mood17			0.805
Supp8			0.723
Func4			0.692
Mood16			0.484

each question loaded onto each of the components is presented in Table 1. Item Rapp10 was removed from the PCA due to its unclear wording, and Supp7 was removed because it had no relation to any of the components.

A brief analysis of the video recordings of each participant revealed that the robot gave 3–9 responses to each participant (with one person receiving 16 responses). Additionally, no one spoke to the robot aside from greeting it, and people did not look at the robot more than twice once they started sorting the medications. With the robot providing relatively few responses and the participants not attempting to respond back to the robot, the level of interactivity was fairly low.

5.2.4 Post-hoc Analysis and Results

During the evaluation, we noticed that some people greeted the robot and others did not. At the beginning of the interaction, the robot introduced itself: "Welcome, my name is Shafer." Some participants chose to respond to this greeting. We reviewed the video of each of the interactions to find that five of the participants responded with some sort of greeting (e.g., saying "hello," waving).

To see if these participants rated the experience with the robot any differently, we compared the scores from the questionnaire for the greeters and non-greeters using an independent sample t test. The non-greeters reported higher means on all of the questionnaire items ($M = 3.84$, $SD = 0.42$) as compared to greeters ($M = 3.69, SD = 0.36$), but the difference was not significant ($p = 0.55$) (see Fig. 7). We used an independent two-sample t test to find that between greeters and non-greeters, there were no significant differences across any of the three components from the PCA:

1. $t(9) = -0.2$, $p = 0.91$, 95% CI[$-1.28, 1.16$]; $d = 0.07$
2. $t(9) = -0.35$, $p = 0.73$, 95% CI[$-0.89, 0.65$]; $d = 0.22$
3. $t(9) = -1.25$, $p = 0.24$, 95% CI[$-1.36, 0.39$]; $d = 0.74$

While none of the components had significant differences, the third component was found to have a large effect size ($d = 0.74$).

5.3 Discussion of the Initial Prototype Evaluation

This prototype helped us explore the design space of a social robot that assists in medication sorting. We discuss here the questions we set forth at the beginning of this prototype, the evaluation, and important next steps.

5.3.1 Design Questions

To guide us in exploring the design space, we established three questions.

How should the robot decide what level of assistance to provide?

We chose to organize the decision process into a binary tree (Fig. 5), where deeper into the tree resulted in a greater level of assistance. However, manually constructing the decision tree was time consuming, and the tree did not provide adequate coverage for all the cases in which assistance was to be provided. In the next prototype, we would need a larger decision tree that could better handle numerous conditions relevant to making the appropriate decision on the type and level of assistance. Additionally, if the complexity of the medication sorting task were to be increased (e.g., more pills, more variability), then the decision tree would need to scale to these additional complexities.

For each level of assistance, what does it mean to give that type of assistance?

We developed concrete examples of assistance at each level and the conditions under which the assistance would be provided, but these examples were not sufficient. Lower level assistance was too repetitive and could become annoying or offensive. Instead of repeating an assistance, more assistance should be provided—giving assistance at a higher level. Higher levels of assistance did not provide enough information about what was wrong or how to fix it. The robot should indicate information about missing pills or which day had too many pills.

And what are the relevant dimensions for evaluating the robot and people's experience with the robot?

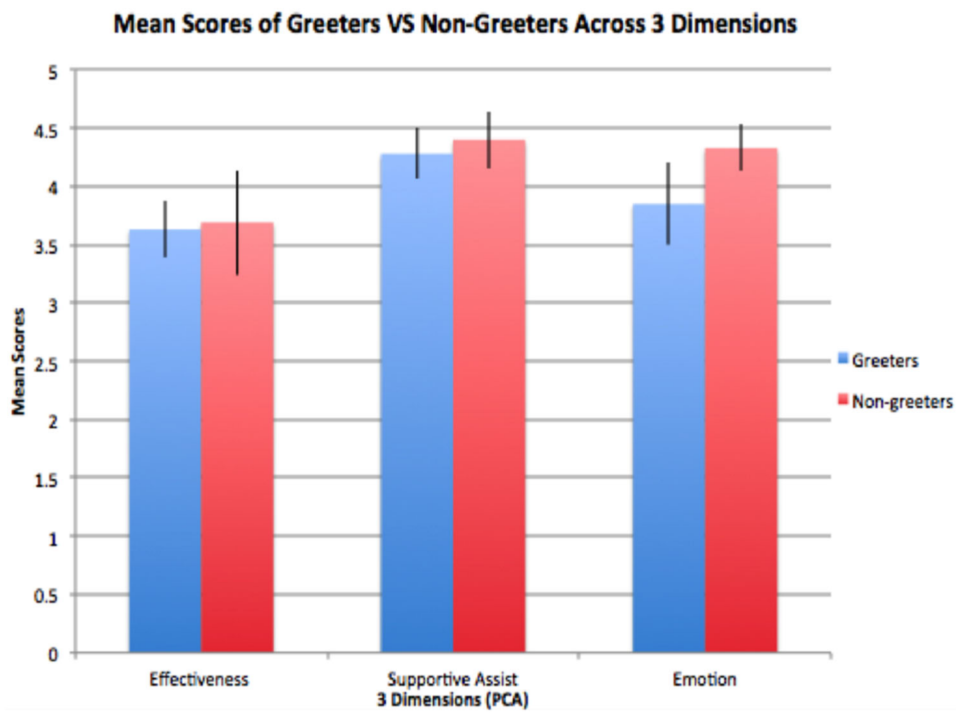

Fig. 7 The non-greeters tended to judge the experience with the robot higher, but not significantly so. In the emotion dimension, the difference had a large effect size (d = 0.74)

In our evaluation, we answered our third question, identifying the relevant components for evaluating the robot. We found that the components that conceptually made the most sense were *effectiveness*, *supportive assistance*, and *emotion*. We continued to use these components in evaluating the next prototype.

5.3.2 Evaluation Highlights

In evaluating our prototype, we gained some evidence that we were accomplishing our design goals to support the autonomy of the person doing the medication sorting. People operating with autonomy likely feel in control and responsible, which was how our participants reported feeling. The highest rated questions had the participants strongly agreeing that they felt in control during the task, understood what was happening, and were responsible for completing the task.

Our evaluation also raised a new question in regards to a possible difference between those who greet the robot and those who do not. Even though there were no significant differences with our small sample size, we found this potential trend to be a little surprising and worth further investigating. One speculation was that perhaps the robot was not sufficiently interactive after the introduction. People could have been disappointed by the robot and its limited abilities after initially seeming very social and interactive. Keeping the robot more interactive and engaging throughout the task (e.g., via responding to natural language) could have altered this trend.

5.3.3 Next Steps

Important to the next prototype, the medication sorting task needed to be more complicated and challenging. The simplicity of the task in this prototype may have led to the high volume of positive responses we received in the evaluation. Ideally, the task would be difficult enough that people would naturally make mistakes and not need to be instructed to make mistakes. Thus, the external validity of our evaluation would be increased by having a more challenging medication sorting task. Additionally, a more complex task would allow us to more fully explore the different conditions under which the robot needs to decide how to assist. In the next prototype we made the task more challenging and enhanced the robot's Memory Management and Assistant component to reason about a more complex task accordingly.

6 The Second Prototype

To design the second prototype, we established a couple of objectives. First, the medication sorting task needed to be more complex in order for us to better explore how the robot decides to assist and the content of the assistance the robot is to provide. As part of this, the task should be complex enough such that the person naturally makes mistakes and the robot is genuinely assisting. Second, the robot needs to interact more with the person doing the medication sorting in order for us to better assess the experience the person has in being assisted by the robot. Lastly, we should consider a lower cost platform than the Nao robot that has nearly identical capabilities. We continued to use the *physical robot*, but we also considered a *virtual robot*, which was a simulation of the Nao robot (Fig. 10).

6.1 Architecture

This prototype incorporated many updates to the architecture. The new set of components and the flow of data are shown in Fig. 8. We introduced a revised medication management and assistance component to handle the more complex task and to better select the appropriate assistance. To support a more interactive robot, we added a speech processing component, added gaze detection to the Vision component, and added a sensory integration component to handle the multiple sources of sensory input.

We also introduced a virtual agent that operated nearly identically to the robotic agent. Controlling the virtual agent did not require any new architectural components because we were using a simulation of the robot; the robot control component worked interchangeably with the physical robot and the virtual robot with no changes. For more details on the hardware setup for the two instantiations of the robot (physical and virtual), see Sect. 6.2.1. One small difference in the platforms is that the virtual agent did not have speech production capabilities. As a result, we needed to add a component to handle the speech production in the case of the virtual agent. In the sections below we describe the new components and those components that underwent significant change.

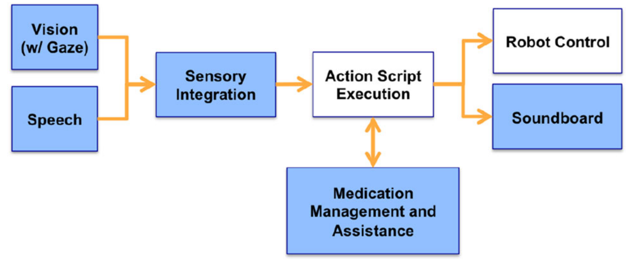

Fig. 8 This prototype added speech, sensory integration, and soundboard components. The action script execution component directly communicated with the robot control to execute all actions, and the medication management and assistance component was only used by the action script execution component to select the appropriate actions. New components or ones with substantial changes are shaded. (Color figure online)

6.1.1 Gaze Detection

We extended the Vision component to detect the direction of the person's gaze, whether they were looking downward at the medication grid or upward at the robot. While the person was sorting medications, they were mostly looking down at the medication grid, but when the person asked the robot a question or was looking for feedback from the robot, they tended to look at the robot.

We used the information about the direction in which the person was looking in two ways. If the person was looking at the robot and the robot was to respond, the robot should look at the person when responding and not continue to look down at the medication grid. The second use of this functionality is that the change in the gaze direction triggered an event. When the person looked up at the robot, the robot decided whether to and how to respond. For example, when the person thought he or she might be done but did not say anything and only looked up at the robot waiting for the next instruction, the robot interpreted the change in the direction of the gaze as a cue to respond and inform the person that the task was not complete.

6.1.2 Speech Recognition

We introduced a new component to provide automatic speech recognition. In this prototype, we were only interested in identifying a few keywords and did not need the speech to be fully recognized or to have the utterance parsed. Identifying the keywords was sufficient to determine whether the person was asking a question, referenced a day of the week, or made a comment about being done. Examples include the following (with keyword emphasized): "*When* do I have physical therapy?", "Is *Monday* ok?", and "I'm *done*". These keywords were later used by the medication management and assistance component to determine the appropriate type of action.

The speech recognition is accomplished using Sphinx4 [27], which was configured to recognize utterances from a user defined grammar in JSGF format. The grammar did not define a strict sentence structure for each utterance and mainly identified the keywords that we needed. For example, the following rule was used to identify question words:

```
<wQuestion> = what | when | where | which| why;
```

6.1.3 Sensory Integration

The purpose of the sensory integration component was to combine inputs from the Vision and Speech components into a single event for the action script execution and medication management and assistance components to handle. This event then included all of the perceptual information necessary for the robot to decide how to respond.

Each time a state change in the Vision system has been detected, a new event was sent to the sensory integration component. As in the initial prototype, a state change consisted of a change in the number of pills at any position in the medication grid. New to this prototype was that changes in the direction of the gaze also caused a new event to be sent.

When the Speech component recognized a new utterance, it was sent to the sensory integration component. Upon receiving a new utterance, the sensory integration component packaged together the utterance along with the last visual perception that was received and passed it along to be processed.

6.1.4 Medication Management and Assistance

When choosing how to respond to a particular event, the action script execution (ASE) component utilized a decision aid to determine how best to respond to each perceived action. This was similar to the architecture in the original prototype but with a couple of small differences. First, the medication management and assistance (MMA) component was only responsible for selecting the action and not for communicating with the robot control component to execute the action on the robot. This change in the communication is shown in the architecture figure (Fig. 8), with the MMA component only connected to the ASE component. Changing the flow of control allowed us to better separate the action selection from the action execution. The MMA component was no longer directly connected to the software to control the robot, making it easier to try the system on a different target hardware. More importantly, using the ASE component to manage the execution of the actions allowed the MMA component to select an assistance that was an *action script*, as opposed to a primitive action that the robot control component would directly recognize. An example of the MMA selecting an action script was when the robot was to look at the person and then speak. This was encoded as a `lookAndSay` action script that had primitive actions of `lookAt` (to look at the person), `sayText` (to produce speech from text), and `lookAt` (to look back down at the medication grid).

Decision Tree Just as in the original MMA component, the reasoning was based on a binary decision tree. However, the new tree was significantly larger (2097 nodes now, as compared to 39 previously) and is procedurally generated instead of being entirely constructed by hand. The structure of the decision tree was also quite different. Previously, the depth in the tree roughly corresponded with the level of assistance selected. Now the level of assistance is handled in a very different manner (see *Level of assistance* in this section) and had no relation to tree depth. The new structure of the decision tree had the upper portion of the tree classifying the event into one of 7 categories:

- Done (pill or all)
- Not started
- Incomplete
- Wrong time
- Wrong day
- Too many
- Wrong time and too many

The *Done* category indicated that either all the pills for a single medication have been completely sorted or all the medications have been completed and the task is done. The *Not started* category was for the beginning of the task. This handled the case where the person was not sure if they should begin or how to begin. *Incomplete* was when there were no misplaced pills but the task was still not complete. Unlike the original MMA component that was primarily triggered on changes in the state of the grid, responses could also be triggered by changes in the direction of the person's gaze or when the person spoke. If there was no change in the state of the grid and the state of the grid was consistent with the goal state of the grid then the event was categorized as *Incomplete*. The *Wrong day* category was for when a pill was placed on a day in which the pill should not to be taken. The *Wrong time* category was for when a pill was placed on the correct day but at the wrong time. A special case of the *Wrong time* category was when not only was a pill at the wrong time but all of the correct times are fully allocated. For example, if a pill was to be taken on Sunday with breakfast, and a person had one pill in the morning cell and another one in the evening, then this was a *Wrong time and too many* event.

Once the category has been determined, then the decisions in the tree were particular to that class of event. The types of decisions included determining which medication had been misplaced, whether the person was looking at the robot or not, and recognizing whether the person had said something and whether the content of the spoken utterance included a particular keyword. For example, for an event categorized as *Incomplete*, the decisions that followed included checking whether a speech utterance was detected, if the the utterance included the word "done", and which medication was incomplete. Then the selected assistance was an action that informed the person about which medication was incomplete.

Level of Assistance Many of the final decisions in the tree pertain to the level of assistance that was necessary. To choose the appropriate level of assistance, we introduced a state variable to track the current level of assistance. Every time a level of assistance greater than the current level was selected, the current level was increased. Similarly, if an action below the current level was selected, the level decreased. This then allowed us to have decisions in the tree that tested against the current level of assistance and selected an action that is at that level or higher.

Action Types There were five types of actions that could be selected: `sayText`, `noOp`, `lookAndSay`, `pointToError`, and `notifyDone`. The first two, `sayText` and noOp, were primitive actions. The noOp action, not doing anything and not responding, was selected frequently, for example, when the person did not need any assistance, or when assistance had been provided recently. This prevented feedback by the robot being generated too frequently.

The other three actions (lookAndSay, pointToError, and notifyDone) were scripts. The lookAndSay action was for when we wanted the robot to look at the person while speaking (as opposed to looking down at the medication grid). This was selected when it had been detected that the person's gaze was directed towards the robot.

The pointToError action is used for the level 4 assistance (gesture) and has the robot point to the location on the grid where a misplaced pill has been detected. The notifyDone action notifies the person that the medication (or all medications) have been successfully sorted, but its more important function is to advance the action script of the overall task so that the robot can either provide new instruction to the person or indicate that the task is complete.

6.1.5 Soundboard

The Soundboard component was only used in the setup for the virtual robot. When the robot was to speak, this component played an audio file that had been mapped to a particular utterance. To ensure that the speech heard by the user was the same for both the physical and virtual robots, we used the physical robot to record each possible utterance to an audio file. A built-in feature of the Nao robot allowed us to redirect the robot's audio output to an audio file, which gave us a clear reproduction of the robot's voice.

6.2 Evaluation

As in the evaluation of the initial prototype, each participant in this study first completed the informed consent form and then performed a medication sorting task with the assistance of the robot. This evaluation had two primary differences from the initial evaluation. First, we wanted to consider a lower cost platform, a virtual robot that was nearly identical to the physical robot. Second, the task was more challenging and more realistic. We also collected some demographics on each person to help us determine what factors may be influencing the human behaviors we would observe.

6.2.1 Hardware Setup

Similar to the setup in the evaluation of the initial prototype, a Nao robot was positioned on the table across from the participant, a medication sorting grid and tray were on the table between the robot and the participant. We constructed a new medication grid and tray (shown in Fig. 9) to have straighter lines (to improve the function of the Vision component) and

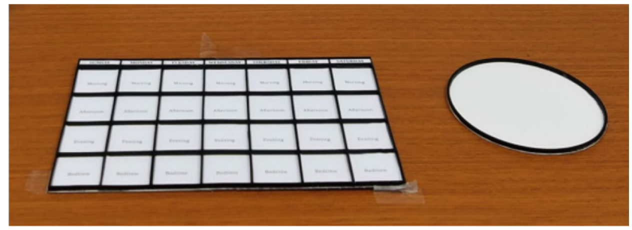

Fig. 9 New medication grid and tray provided clearer lines and divisions, which improved the accuracy of the Vision component

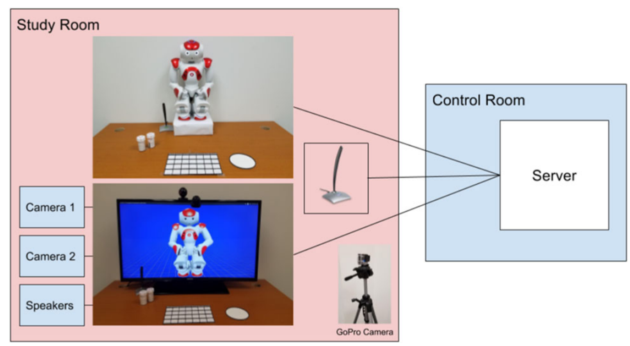

Fig. 10 The robots in both of the physical and virtual setups were controlled by a server running in another room. The virtual robot required additional cameras and speakers to match the hardware provided by the physical robot

to have more raised lines (to prevent pills from accidentally rolling into an adjacent space).

The Nao robot was used for both the physical and virtual conditions. See Fig. 10 for the two different setups. When the participant entered the room it was in a crouched position and was looking downward. The virtual robot was displayed on a 36" display using the Robot View of Choregraphe at full screen. Two webcams were mounted to the top of the display, one pointing downward and one pointing forward (toward the participant). The two built-in cameras in the head of the Nao were used for the physical robot.

To increase the similarity between the conditions, we closely match the size and position of the two robots. Table 2 reports the size and position of each robot. In the physical case, the robot was positioned on a small platform to approximate the position of the virtual robot as it was displayed on the screen.

6.2.2 Evaluation Setup

Participants Students (N = 20) from Tufts University were recruited by word of mouth, signs, and web postings to participate in a social robot study. Half were undergraduate students, and the other half were graduate students. Half of the participants were female. Each participant was randomly assigned to one of two conditions: physical robot or virtual

Table 2 The physical and virtual robots had similar dimensions and were positioned similarly on the table

	Physical (cm)	Virtual (cm)
Height of robot	46	45.5
Width of robot at head	13.5	13
Width of robot at shoulders	26.5	22.5
Bottom of robot from tabletop	8.5	10
Distance of robot from front of table	42	41
Distance of microphone from center	14	27

The microphone in the virtual robot was positioned farther to the left to avoid the display stand but as close to the participant as possible

robot. We continued to use students in our evaluation to facilitate the rapid prototyping.

Methods and Design Each participant interacted with the robot to complete a medication sorting task and then completed a 20-item questionnaire similar to the one used in evaluating the first prototype. The wording on question 10 was modified to remove the negative framing. All other items remained the same. After the questionnaire, each participant answered demographics questions to collect information about major, undergraduate/graduate student status, and gender. Finally, there were three questions asking the participant if they have seen movies with robots, where they have seen robots in their lives, and have they interacted with a robot before.

The containers of simulated medicines were actual pill bottles. On each bottle we affixed a label that named that medication and provided instructions on how the medication was to be taken. The instructions for one was simple: take one with breakfast. The other said the following: "Take 2 with meal. Enhances motor movement. Cannot take more than two pills every other day."

Before the task started, the research assistant oriented the participant to the task by reading a script. Included in the script were details about the person's schedule that would affect when the medications were to be taken. The participant was told that the schedule could not be written down. The intent was to create a small load on working memory, hopefully creating enough of a challenge that participants would naturally make mistakes and need assistance from the robot. Additionally, this could partially simulate the impact on cognitive function that would be experienced by older adults that could eventually be using our system.

6.2.3 Analysis

Two participants in each condition were not able to complete the task. In each case there were technical difficulties with the robot which likely affected the person's ability to complete the task. Since these participants were not fully able to experience the system, they were removed from the analysis. Of the 8 participants remaining in the physical robot condition, 4 were graduate students and 4 were undergraduates, and 4 were female and 4 were male. In the virtual robot condition, 4 were graduate students and 4 were undergraduates, and 5 were female and 3 were male.

The video of each participant was reviewed to determine start and end times of the task and whether the participant greeted the robot after it introduced itself. One participant did not consent to being video recorded, and that person was not able to complete the task due to a technical failure with the robot.

We compared the questionnaire answers and time to complete the task across conditions using unpaired t tests. We also compared questionnaire results using the 3 components we found in the evaluation of the initial prototype. Using an unpaired t test, we compared the first prototype and this prototype (collapsing across both virtual and physical) using these 3 components. Finally, we reviewed the video recordings to determine who greeted the robot and who did not. Two-way ANOVAs were used to compare greeters and non-greeters of the virtual robot and the physical robot. Additionally, we calculated a Cohen's d effect size for each comparison.

6.2.4 Evaluation Results

In this section we first discuss the results of the individual questionnaire answers. Then we address the results relating to grouping the questionnaire answers according to their components. Finally, we compare the results of greeters and non-greeters.

Comparing the questionnaire answers for the physical robot and virtual robot conditions showed no significant differences on any of the questions. See Fig. 11a for a comparison of each question across the two conditions. In comparing the time it took to complete the task, those being assisted by the virtual robot finished the task more quickly (Virtual: $M = 3.07$, $SD = 1.02$; Physical: $M = 4.93$, $SD = 3.04$), but the difference was not significant ($t(14) = 1.64$, $p = 0.12$). We also evaluated the effect sizes in order to determine the magnitude of the effect within this sample, despite there not being enough evidence for generalization to the broader pop-

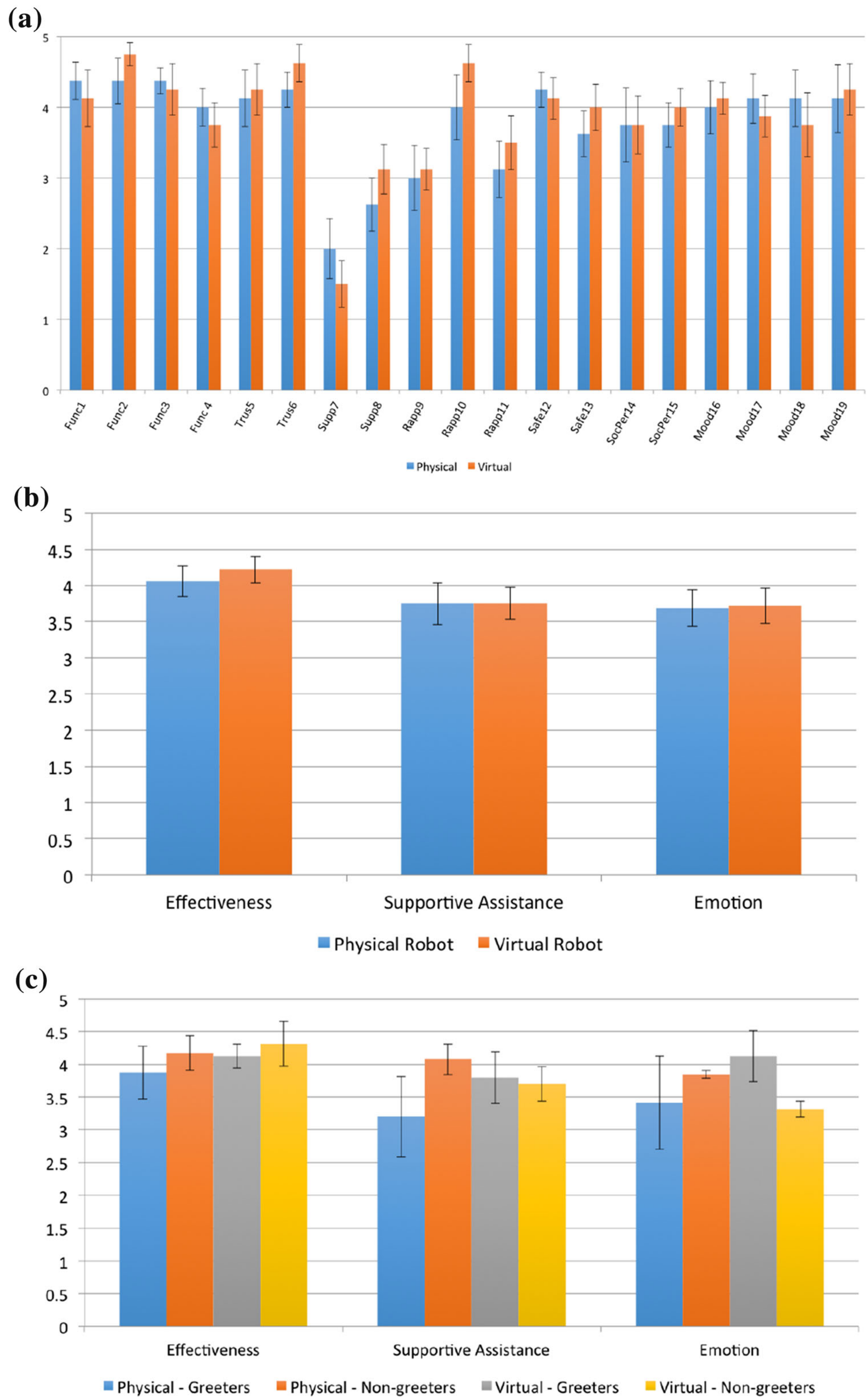

Fig. 11 Results of evaluation of second prototype. Each figure shows mean results for items in a questionnaire answered by a participant that interacted with the physical robot or the virtual robot. Error bars show standard error. **a** Shows all questionnaire items individually, comparing answers for the physical and virtual robots. **b** Comparing questionnaire answers aggregated into the dimensions of effectiveness, supportive assistance, and emotion. **c** Comparing greeters and non-greeters for each of the dimensions

ulation. The effect size for time physical versus virtual in time to complete the task was large ($d = 0.82$).

We calculated component scores of the items for *effectiveness*, *supportive assistance* and *emotion* that were developed in the evaluation of the first prototype. In comparing the component scores of the first and second prototype, we found that the *effectiveness* of the second prototype was higher but not significantly ($t(25) = 1.78$, $p = 0.09$), the *supportive assistance* for the second prototype was significantly lower ($t(25) = 2.59$, $p = 0.02$), and the *emotion* for the second prototype was lower but not significantly ($t(25) = 1.67$, $p = 0.11$). The lower scores for the *supportive assistance* was probably a result of the task used in evaluating the second prototype being more difficult.

When comparing the component scores of the physical and virtual robot for the second prototype, the results showed clearly that participants perceived a strong similarity between the physical and virtual robots (see Fig. 11b). However, difference in sample effect sizes increased in magnitude when comparing participants who greeted the robot and those who did not. Three of the 9 participants greeted the physical robot, and 4 of the 9 participants greeted the virtual robot. Figure 11c shows the same three components but splits the greeters from the non-greeters. In the first component, *effectiveness*, physical and virtual and greeters and non-greeters all reported very similar answers with no significant main effect for robot type ($F(1, 12) = 0.28$, $p = 0.60$) or greeting ($F(1, 12) = 0.66$, $p = 0.43$) and no interaction ($F(1, 12) = 0.04$, $p = 0.85$). In the second component, *supportive assistance*, there was no difference between the virtual robot and the physical robot ($F(1, 12) = 0.0$, $p = 1.0$), no effect for whether the person greeted the robot ($F(1, 12) = 1.08$, $p = 0.32$), and no interaction ($F(1, 12) = 1.86$, $p = 0.20$). Even though there was no statical significance for the interaction, there was a large effect size for greeting the physical robot ($d = 1.24$) but not for greeting the virtual robot ($d = 0.14$). The *emotion* component also had some possible differences, but mostly for the virtual robot. There was no effect for robot type ($F(1, 12) = 0.01$, $p = 0.93$) or for greeting the robot ($F(1, 12) = 0.40$, $p = 0.54$), but an interaction was trending towards significance ($F(1, 12) = 3.53$, $p = 0.08$). In this interaction, it is important to note that the direction of the effect was opposite for the virtual robot and the physical robot. Greeting the physical robot was related to a lower rating, and had a small to medium effect size ($d = 0.66$). Greeting the virtual robot was related to a higher rating and had a large effect size ($d = 1.25$). This suggests that the social aspects contributing to the emotion of the experience differ for a virtual robot and a physical robot, perhaps due to people being more familiar with interacting with a screen as opposed to a physical robot.

6.3 Discussion

The results suggested that people who greeted the robot experienced the interaction differently, and that the nature of this difference varied depending on whether the person interacted with the physical or virtual robot. For the physical robot, greeters of the robot rated the *supportive assistance* of the experience lower than non-greeters. This effect was not significant with our small sample size, but the large effect size ($d = 1.24$) suggests that it is feasible that greeting the robot was related to how the physical robot's assistance was perceived. Similarly, the greeters of the physical robot provided lower ratings for the *emotion* component. In addition to there being a medium effect size ($d = 0.66$), it is interesting to note that we saw a similar pattern in our initial evaluation. This provides more evidence that greeting a physical robot had an effect on our participants, but the lack of significance due to a small sample size prevents us from drawing any conclusions about the population of users of physical robots. Greeters of the virtual robot appeared to have had a different experience in regards to the *emotion* component. Greeters of the virtual robot rated the *emotion* of the experience *higher*, and the analysis gave a large effect size ($d = 1.25$). This result was not significant, but it is feasible that greeting the virtual robot was related to the higher rating of the *emotion* component.

If, however, the differences in how the greeters and non-greeters experienced the two robots are negligible or inconsequential, then the virtual robot may be an ideal low-cost alternative. The virtual robot would be substantially less expensive, which would improve the likelihood of older adults being able to afford the assistive technology. However, before we can commit to the virtual robot, we need to consider some of the limitations of the virtual robot.

One limitation of a virtual robot is that it would not be able to provide physical support (e.g., demonstrate how to place a pill or guiding a person's arm to help initiate movement). In our prototypes, the physical robot was also designed to not provide physical support, but this was a limitation of our design goal (to only provide cognitive and social support) and not a limitation inherent in the hardware platform. If we were to use a physical robot capable of manipulating pills or helping move a person's arm, then the physical robot would be able to provide assistance at levels greater than our current design. However, the potential benefits of the robot providing physical support may not outweigh the ethical concerns around physical contact with a person [17].

7 General Discussion

We have presented a human-centric approach to designing a social robot to assist with medication sorting in a manner that supports the autonomy and dignity of the person. Novel in our approach was the use of a standardized taxonomy defining levels of assistance. The robot selected the appropriate level of assistance by matching the need of the individual to the assistance. In exploring designs for how the robot was to make this selection in a medication sorting task, we identified relevant features of the task that the robot can use as the basis of the decision, but our solutions thus far were specific to medication sorting. We are working to generalize this approach so that we can easily adapt the robot to assist in other activities of daily living. While we are working to generalize this approach, we still need to resolve some other issues.

One such issue is that it remains unclear as to whether a physical or virtual robot was better for assisting with medication sorting. Based on previous studies on other tasks [4,26,29], we expected the physical robot to be rated better. However, we were surprised to find this was not the case. In evaluating our second prototype, participants rated the two platforms similarly. If there are few differences between the platforms, and the virtual robot performs equally as well as its physical counterpart, then the virtual robot could be an ideal low-cost solution that older adults could afford. Based on prior results with physically present robots [30,48,53], we expected the physical robot to improve the interaction, but future work will need to further investigate the potential role of other options, such as smartphone apps. However, many of these alternatives will have some inherent limitations, namely an inability to physically interact with the environment. While we focused our design on cognitive and social support, at least a physical robot has the potential to provide some physical support. This physical support is essential in providing higher levels of assistance, and as we continue to investigate the effectiveness of a robot in assisting with medication sorting, we need to at least consider these higher levels of assistance.

One difference between physical and virtual robots that we have found was that those who voluntarily greeted the robot rated the experience with the robot differently than those who chose not to greet the robot. The difference was most apparent in regards to the *emotion* component, where greeters of the physical robot rated the experience lower than non-greeters and greeters of the virtual robot rated the experience higher. It should be noted that the effect of greeters rating the physical robot lower in regards to emotion that was observed in the first prototype was duplicated in the second prototype. Even though we did not have statistical significance in either case, we did have a medium or large effect size in both evaluations. In speculating why greeters and non-greeters would rate the experience differently, we suspect that it is related to how much the person perceives the robot to be a social agent. A social robot has the potential to provide benefits such as decreased loneliness [25] and more positive mood [51] in addition to assisting in important activities, and the physical embodiment may facilitate some social interactions, as was seen with the Paro robot [50]. From our evaluations and the examples given here, it was apparent that there were social and emotional aspects to the robot that need to be considered in future designs.

As we continue to investigate the designs for a social robot to assist in medication sorting, we need to include older adults in our evaluations. Using students has allowed us to rapidly prototype and evaluate different designs, but the end goal is still to improve the quality of life of older adults. We are particularly interested in how a social robot could assist a person with Parkinson's disease (PD), which is a neurological disorder affecting the dopamine pathways of the brain. Physical symptoms like tremors, fine motor control, and slowness of movement can make medication sorting difficult. Cognitive impairment, another common symptom, can add even greater challenges to the task. Also, unlike many disorders that can simply use a set schedule for medications every day, a person with PD often requires a flexible schedule that can adapt to the life events of the person. It is common for a person with PD to have to take a dopamine agonist or L-DOPA before activity like a dance or exercise class or doctor appointment. We envision a social robot that can have a dialogue with the person to collaboratively determine the best time to take the medications that fits the individual's schedule and that complies with the constraints of the prescription. The conversation the robot has with the person should help the person feel and be more involved in health-related decisions. Additionally, the robot being able to adapt to the person and the person's schedule should help the person feel validated as an individual whose distinctive preferences and qualities matter. Overall, this supports one of our key design principles of having the robot help preserve the person's dignity.

8 Conclusion

We have demonstrated a robot the provides cognitive and social support to a person sorting medications. Instead of providing direct physical support to the person, the robot enabled the person to make the correct actions to successfully sort the medications. Our novel approach to selecting the appropriate assistance utilized a taxonomy defining increasing levels of assistance. The taxonomy, which is used in practice by occupational therapists, informed the decision process by which the robot provided the amount of assistance that met the need of the person.

We developed two prototypes of a robot designed to assist with medication sorting. Upon evaluating the first prototype, we identified three components used to evaluate the interaction with the robot: *effectiveness*, *supportive assistance*, and *emotion*. The second prototype built upon the first and add more social cues (i.e., gaze and speech) for responding to the person sorting the pills, and we increased the complexity of the medication sorting task to improve the ecological validity of our design. In comparison with the first prototype, this robot was rated slightly higher for *effectiveness* but lower in *supportive assistance*. Both of these effects may have been a result of the increased task complexity, but future work would be necessary to determine if it was the task complexity or some other factor that caused these effects.

For the second prototype, we also considered a virtual robot, a robot that was nearly identical to the physical robot but simulated in a graphical environment and displayed on a screen. In general, each of the embodiments of the robot (physical and virtual) were evaluated similarly, which is in contrast to prior work that has shown preferences for the physical robot over the virtual [4,34]. More work will be needed to understand why the physical robot may be preferred in some cases but in other cases neither the physical robot nor the virtual robot is preferred.

In evaluating both of the prototypes, we discovered an interesting factor affecting the experience of being assisted by the robot. For the first prototype, we found that those who voluntarily greeted the robot rated the *emotion* of the experience lower. This effect was replicated in evaluating the second prototype, but only for the physical robot. Instead, the greeters of the virtual robot rated the emotion of the experience higher. Furthermore, we found a small effect on the *supportive assistance* rating, with the greeters of the physical robot rating the experience lower than the non-greeters. It is interesting to note that these effects were observed even though greeting the robot should be orthogonal to the task of sorting medications. This has implications for any task in which the robot is providing assistance, as social aspects of the interaction that are apparently unrelated to the task are likely to still affect the experience.

As we proceed in designing robots to assist older adults, we need to continue to take a human-centric approach. We cannot simply look at how well the robot can assist on a task, but we need to consider how the robot is perceived and the effects the robot may have on the person being assisted. We believe that by matching the need of the person to the assistance provided, the robot can allow the person to feel in control and contribute to the person's sense of autonomy.

Acknowledgements This work was funded by NSF (Grant No. IIS-1316809).

Compliance with Ethical Standards

Conflict of interest All authors declare that they have no conflict of interest.

Appendix

The following 19 questions were provided to each participant. Question 10 had a negative working in the first evaluation and was updated for the second evaluation (working from 2nd presented here).

For each question, a 5-point Likert-scale was provided, ranging from strongly disagree to strongly agree.

1. The robot is able to provide you with assistance in the task.
2. The assistance the robot provides is correct.
3. I am able to complete the task more efficiently with the assistance of the robot.
4. When the robot corrects me I feel included to follows its instructions.
5. I trust the robot to (correctly) provide assistance.
6. I expect the robot to act in a consistent and predictable manner.
7. The robot is able to provide physical support.
8. The robot is able to provide emotional support.
9. The robot paid attention to me.
10. The robot used action and words that made sense to me.
11. The robot helped me understand how to complete the task.
12. The robot acted in a manner that ensured my safety.
13. The robot is able to warn me of potentially unsafe medication administration.
14. My family would approve of the way the robot assisted me.
15. My care providers would approve of the way the robot assisted me.
16. I felt pleasant during the task.
17. I felt in control of what was happening during the task.
18. I felt I understood what was happening during the task.
19. I felt responsible for completing the task.

    In both evaluations, we also included the following 20th question:

20. Do you have anything else you want to tell us about the robot or this task?

References

1. American Journal of Occupational Therapy: Occupational Therapy Practice Framework: Domain and Process (3rd edn) (2014)
2. Association AOT et al (2015) Occupational therapy code of ethics (2015). Am J Occup Ther 69:S17–S26
3. Bainbridge JL, Ruscin JM (2009) Challenges of treatment adherence in older patients with parkinson's disease. Drugs Aging 26(2):145–155
4. Bainbridge WA, Hart JW, Kim ES, Scassellati B (2011) The benefits of interactions with physically present robots over video-displayed agents. Int J Soc Robot cs 3(1):41–52
5. Banks MR, Willoughby LM, Banks WA (2008) Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J Am Med Dir Assoc 9(3):173–177
6. Baum E (2006) Pill dispenser. US patent 7,100,793
7. Briggs P, Scheutz M, Tickle-Degnen L (2015) Are robots ready for administering health status surveys? first results from an HRI study with subjects with Parkinson's disease. In: Proceedings of the 10 annual ACM/IEEE international conference on human–robot interaction. ACM, Portland, pp 327–334
8. Broekens J, Heerink M, Rosendal H (2009) Assistive social robots in elderly care: a review. Gerontechnology 8(2):94–103

9. Burnier M (2006) Medication adherence and persistence as the cornerstone of effective antihypertensive therapy. Am J Hypertens 19(11):1190–1196
10. Carlson MC, Fried LP, Xue QL, Tekwe C, Brandt J (2005) Validation of the hopkins medication schedule to identify difficulties in taking medications. J Gerontol Ser A: Biol Sci Med Sci 60(2):217–223
11. Chan M, Nicklason F, Vial J (2001) Adverse drug events as a cause of hospital admission in the elderly. Int Med J 31(4):199–205
12. Datta C, Yang HY, Tiwari P, Kuo IH, MacDonald BA (2011) End user programming to enable closed-loop medication management using a healthcare robot. Social Science, New York
13. Dayer L, Heldenbrand S, Anderson P, Gubbins PO, Martin BC (2013) Smartphone medication adherence apps: potential benefits to patients and providers. J Am Pharm Assoc 53(2):172–181
14. Doggrell SA (2010) Adherence to medicines in the older-aged with chronic conditions. Drugs Aging 27(3):239–254
15. Ernst FR, Grizzle AJ (2001) Drug-related morbidity and mortality: updating the cost-of-illness model. J Am Pharm Assoc 41(2):192–199 (1996)
16. Feil-Seifer D, Matarić MJ (2005) Defining socially assistive robotics. In: ICORR 2005. 9th international conference on rehabilitation robotics, 2005, pp 465–468
17. Feil-seifer D, Matarić MJ (2011) Ethical principles for socially assistive robotics. Robot Autom Mag 18(1):24–31
18. Graf B, Hans M, Schraft RD (2004) Care-o-bot II—development of a next generation robotic home assistant. Auton Robots 16(2):193–205
19. Hayes TL, Larimer N, Adami A, Kaye JA (2009) Medication adherence in healthy elders: small cognitive changes make a big difference. J Aging Health 21:567–580
20. Haynes RB, Ackloo E, Sahota N, McDonald HP, Yao X (2008) Interventions for enhancing medication adherence. Cochrane Database Syst Rev 16(2):CD000011
21. Hebesberger D, Dondrup C, Koertner T, Gisinger C, Pripfl J (2016) Lessons learned from the deployment of a long-term autonomous robot as companion in physical therapy for older adults with dementia: a mixed methods study. In: The 11th ACM/IEEE international conference on human robot interaction. IEEE Press, pp 27–34
22. Heerink M, Kröse B, Wielinga B, Evers V (2006) Human–robot user studies in eldercare: lessons learned. In: International conference on smart homes and health telematics, pp 31–38
23. Hemmesch AR (2011) The stigmatizing effects of facial masking and abnormal bodily movement on older adults' first impressions of individuals with Parkinson's disease. Brandeis University
24. Johnson JA, Bootman JL (1995) Drug-related morbidity and mortality: a cost-of-illness model. Arch Intern Med 155(18):1949–1956
25. Kanamori M, Suzuki M, Tanaka M (2002) Maintenance and improvement of quality of life among elderly patients using a pet-type robot. Jpn J Geriatr 39(2):214–218
26. Kidd CD, Breazeal C (2004) Effect of a robot on user perceptions. In: Proceedings. 2004 IEEE/RSJ international conference on intelligent robots and systems, 2004 (IROS 2004), vol 4. IEEE, pp 3559–3564
27. Lamere P, Kwok P, Gouvea E, Raj B, Singh R, Walker W, Warmuth M, Wolf P (2003) The CMU sphinx-4 speech recognition system. In: IEEE international conference on acoustics, speech and signal processing (ICASSP 2003), Hong Kong, vol 1, pp 2–5, Citeseer
28. Lewis A (1997) Non-compliance: a $100 billion problem. Remington Rep 5(4):14–15
29. Li J (2015) The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum Comput Stud 77:23–37
30. Li J (2015) The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum Comput Stud 77(January 2015):23–27
31. MacLaughlin EJ, Raehl CL, Treadway AK, Sterling TL, Zoller DP, Bond CA (2005) Assessing medication adherence in the elderly: which tools to use in clinical practice? Drugs Aging 22(3):231–255
32. Manning KJ, Clarke C, Lorry A, Weintraub D, Wilkinson JR, Duda JE, Moberg PJ (2012) Medication management and neuropsychological performance in parkinson's disease. Clin Neuropsycholog 26(1):45–58
33. Marcum ZA, Sevick MA, Handler SM (2013) Medication non-adherence: a diagnosable and treatable medical condition. JAMA 309(20):2105–2106
34. Matarić MJ, Eriksson J, Feil-Seifer DJ, Winstein CJ (2007) Socially assistive robotics for post-stroke rehabilitation. J NeuroEng Rehab 4:5
35. Muir BM, Moray N (1996) Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics 39(3):429–460
36. Osterberg L, Blaschke T, Koop CE (2005) Adherence to Medication. N Engl J Med 353:487–497. https://doi.org/10.1056/NEJMra050100
37. Pollack M, Brown L, Colbry D (2002) Pearl: a mobile robotic assistant for the elderly. In: AAAI workshop on automation as eldercare, pp 85–91
38. Prakash A, Beer JM, Deyle T, Smarr CA, Chen TL, Mitzner TL, Kemp CC, Rogers WA (2013) Older adults' medication management in the home: How can robots help? In: Proceedings of the 8th international conference on human–robot interaction, pp 283–290
39. Rogers J, Holm M (1994) Performance assessment of self-care skills (pass-home) version 3.1. University of Pittsburgh, Pittsburgh [unpublished assessment tool]
40. Sahai A, Breede SK, Topliffe RO, Topliffe DA (1999) Medication dispensing system. US patent 5,971,594
41. Schell BA, Gillen G, Scaffa M, Cohn ES (2013) Willard and Spackman's occupational therapy. Lippincott Williams and Wilkins, Philadelphia
42. Scheutz M, Schermerhorn P, Kramer J, Anderson D (2007) First steps toward natural human-like HRI. Auton Robots 22(4):411–423
43. Smarr CA, Prakash A, Beer JM, Mitzner TL, Kemp CC, Rogers WA (2012) Older adults' preferences for and acceptance of robot assistance for everyday living tasks. In: Proceedings of the human factors and ergonomics society 56th annual meeting, pp 153–157
44. Smith CA, Ellsworth PC (1985) Patterns of cognitive appraisal in emotion. J Pers Soc Psychol 48(4):813–838
45. Sokol MC, Mcguigan KA, Verbrugge RR, Epstein RS (2005) Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 43(6):521–530
46. Stilley CS, Bender CM, Dunbar-Jacob J, Sereika S, Ryan CM (2010) The impact of cognitive function on medication management: three studies. Health Psychol 29(1):50
47. Tamura T, Yonemitsu S, Itoh A, Oikawa D, Kawakami A, Higashi Y, Fujimooto T, Nakajima K (2004) Is an entertainment robot useful in the care of elderly people with severe dementia? J Gerontol Ser A: Biol Sci Med Sci 59(1):M83–M85
48. Tapus A (2009) The role of physical embodiment of a therapist robot for individuals with cognitive impairments. In: Virtual rehabilitation international conference, 2009
49. Wada K, Shibata T (2006) Robot therapy in a care house—its sociopsychological and physiological effects on the residents. In: Proceedings of the 2006 IEEE international conference on robotics and automation, pp 3966–3971
50. Wada K, Shibata T (2007) Robot therapy in a care house—change of relationship among the residents and seal robot during a 2-month long study. IEEE Trans Robot 23(5):972–980

51. Wada K, Shibata T, Saito T, Tanie K (2002) Robot assisted activity for elderly people and nurses at a day service center. In: Proceedings of the 2002 IEEE international conference on robotics and automation, Washington, DC, pp 1416–1421
52. Wade R, Clancey B, Michaeli J (2016) Improvement in antihypertensive and cholesterol-lowering medication persistence using a mobile technology application. Value Health 19(3):A306
53. Wainer J, Feil-Seifer DJ, Shell DA, Matarić MJ (2006) The role of physical embodiment in human–robot interaction. In: Proceedings—IEEE international workshop on robot and human interactive communication, pp 117–122
54. Warner I (1998) Telemedicine in home health care: the current status of practice. Home Health Care Manag Pract 10(2):62–68
55. Wilcock W, Townsend E (2014) Occupational justice. Willard and Spackman's Occupational Therapy, pp 541–552

Jason R. Wilson is a Postdoctoral Fellow in the Department of Electrical Engineering and Computer Science at Northwestern University. He earned a Joint Ph.D. in Cognitive Science and Computer Science from Tufts University in 2017. His current research focuses on human-centric applications of artificial intelligence, including socially assistive robots and affective computing.

Nah Young Lee is a Masters student in Occupational Therapy at Tufts University. She received her ME.d. in Human Development and Education from Harvard Graduate School of Education in 2013. Her research interest focuses on social and cognitive functioning and the development of a community-based program for young children with the neurological disease. She is currently working on her fieldwork affiliation at Cambridge Public School.

Annie Saechao is a MSOT candidate in the Department of Occupational Therapy at Tufts University. She has earned a B.S. in Human Development and a minor in Psychology at the University of California, Davis. She has a special interest in assistive technology and how to integrate it within her OT practice.

Linda Tickle-Degnen is Professor in the Department of Occupational Therapy at Tufts University, with a secondary appointment in the Department of Psychology. She received a PhD in experimental social psychology from Harvard University and an MS in occupational therapy from University of Southern California. She directs the Health Quality of Life Lab, which studies health and well-being outcomes of facial and bodily movement, social and cognitive functioning, cross-cultural stigmatization and self-management activity for community living with Parkinson's disease and other chronic and disabling conditions.

Matthias Scheutz is a Professor in Cognitive and Computer Science in the Department of Computer Science at Tufts University. He earned a Ph.D. in Philosophy from the University of Vienna in 1995 and a Joint Ph.D. in Cognitive Science and Computer Science from Indiana University Bloomington in 1999. He has more than 300 peer-reviewed publications in artificial intelligence, natural language processing, cognitive modeling, robotics, and human–robot interaction. His current research focuses on complex cognitive robots with natural language and machine learning capabilities.