Are Preferences Useful for Better Assistance?: A Physically Assistive Robotics user study

GERARD CANAL, Institut de Robòtica i Informàtica Industrial, CSIC-UPC and Department of Informatics, King’s College London
CARME TORRAS and GUILLEM ALENYÀ, Institut de Robòtica i Informàtica Industrial, CSIC-UPC

Assistive Robots have an inherent need of adapting to the user they are assisting. This is crucial for the correct development of the task, user safety, and comfort. However, adaptation can be performed in several manners. We believe user preferences are key to this adaptation. In this paper, we evaluate the use of preferences for Physically Assistive Robotics tasks in a Human-Robot Interaction user evaluation. Three assistive tasks have been implemented consisting of assisted feeding, shoe-fitting, and jacket dressing, where the robot performs each task in a different manner based on user preferences. We assess the ability of the users to determine which execution of the task used their chosen preferences (if any). The obtained results show that most of the users were able to successfully guess the cases where their preferences were used even when they had not seen the task before. We also observe that their satisfaction with the task increases when the chosen preferences are employed. Finally, we also analyze the user’s opinions regarding assistive tasks and preferences, showing promising expectations as to the benefits of adapting the robot behavior to the user through preferences.

CCS Concepts: • Computer systems organization → Robotics; • Human-centered computing → User models; Empirical studies in interaction design; • Computing methodologies → Planning and scheduling.

Additional Key Words and Phrases: HRI, Physically Assistive Robotics, User Preferences, Assistive Feeding, Assistive dressing, Shoe fitting

ACM Reference Format:

1 INTRODUCTION

Physically Assistive Robots (PAR) will be key in the next health care revolution. In the current ageing world, the lack of care professionals such as nurses and caregivers will deem this kind of devices necessary for many individuals to live a dignified life. However, research on assistive robotics poses several challenges, both technical and ethical, given the closeness of the robots to the human users and the many safety concerns this can raise [35].

While we believe these assistive robots cannot (and probably should not) supply the warmth a human caregiver provides, we look at them as smart appliances that provide effective assistance in

Authors’ addresses: Gerard Canal, Institut de Robòtica i Informàtica Industrial, CSIC-UPC, gcanal@iri.upc.edu, Department of Informatics, King’s College London, Bush House, 30 Aldwych, London, WC2B 4BG, United Kingdom; Carme Torras, torras@iri.upc.edu; Guillem Alenyà, galenya@iri.upc.edu, Institut de Robòtica i Informàtica Industrial, CSIC-UPC, C/Llorens i Artigas 4-6, Barcelona, Spain.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
© 2021 Copyright held by the owner/author(s).
2573-9522/2021/09-ART39
https://doi.org/10.1145/3472208

a highly autonomous manner. To successfully assist people with different degrees of dependency, robots must perform the tasks as good as human carers would do. Achieving this will not only increase the robot acceptance rates but also empower users by allowing them to perform Activities of Daily Living (ADLs) autonomously, allowing the human carer to do tasks with a better added-value [1, 12, 23]. However, one of the main elements that humans do when assisting someone is to adapt to their specific needs and preferences. Because no person is identical and each of us has their own tastes, what may work for a user may be disliked for another one. Thus, an increase in task variability implies the need for robot adaptation.

A clear example of this can be seen in Figure 1. As can be observed in the pictures, extracted from Certified Nursing Assistant (CNA) training videos, there are many different ways of doing a simple task such as feeding a person. The strategy will depend on the condition and capabilities of the users, but their preferences will also determine the strategy and play an important role in the success of the task.

In previous works, we have defined a personalization framework and also the potential preferences that can be considered in Physically Assistive Robotics tasks [6, 7], and we studied how to integrate such preferences autonomously using AI Planners [8].

In this paper, we will use our previous work to evaluate if users can identify executions of the assistive tasks in which the robot was using their chosen preferences, without seeing previous examples of the task. Therefore, we will use the notion of preferences for assistive robotics as described in [7], where we defined a hierarchical taxonomy of preferences for guiding robot action selection as well as tuning execution to best meet the user needs and maximize their comfort.

---

1 Extracted from: youtu.be/aCIKmu4jFWg
2 Extracted from: youtu.be/E1vd1O-LgWI
3 Extracted from: youtu.be/j_D8g2ngVWs
4 Extracted from: youtu.be/XvM76BvAoM
5 Extracted from: youtu.be/m4o5u7x4qys
Details on the preferences used in the current work will be provided in Section 3.2, consisting of both task-related preferences guiding action selection and task-independent ones for tuning the robot velocity and communication verbosity. We also consider whether the tasks fulfilling the user’s preferences are found more pleasant and whether the able users see potential benefits for dependent users. Finally, the Almere model [22] has been used to evaluate robot acceptance. For this purpose, three typical assistive tasks have been considered: assisted feeding, shoe-fitting, and jacket dressing. These tasks can support the autonomy of potential (dependent) users by allowing them to perform the tasks by themselves, without the help of a human carer. Two robot arms have been used to assist the users in performing the tasks in a fully autonomous manner. The users have experienced each task two times, answering a questionnaire of their experience after each task. The following sections will detail the used experimental methodology, evaluation measures, and experiment results.

2 RELATED WORK

Physically Assistive Robots require a proper interaction with the assisted user to complete the task successfully. Given the importance of such kind of applications, many works intend to solve some or part of the assistive tasks. Silva et al. [32] propose a modular robotic arm for assisted feeding. Vila et al. [38] assume that impacts may occur accidentally and analyze impact forces and safety measures for robot-assisted feeding. A taxonomy of manipulation strategies for feeding is presented by Bhattacharjee et al. [3]. To do so, they created a dataset of food manipulation. Following up, bite acquisition is further studied by Gallenberger et al. [19], where adaptive strategies select the manipulation primitive to use with each food item. A user study with 25 participants was performed in which users had to determine how easy was to bite off the fork when the robot presented the food.

Yamazaki et al. [40] describe bottom dressing assistance by a robot, where it pulls the clothing along the user’s legs. The authors integrate visual failure detection and a success study is performed with different clothing types. A quantitative analysis of dressing dynamics is performed by Chance et al. [11], where they analyze combinations of user pose, clothing types, and different sensors for detecting errors such as cloth snagging. A hospital gown is dressed by Erickson et al. [18], where a deep recurrent model is used to predict the garment forces applied to the user. This is further applied for improving the robot’s control, using only haptic and kinematic data from the robot’s end-effector. Zhang et al. [42] propose a tracking method based on Bayesian Networks in latent spaces to fuse robot pose and force for camera-less estimation of user postures while dressing. They applied it to the dressing of a sleeveless jacket. The system is evaluated with able-bodied users who had some of the movements restricted by braces to simulate user limitations. These works show that physically assistive robots have received increasing interest from the research community. Still, while some of them focus on adaptivity to the user, none has yet performed a user evaluation on the perception of behavior adaptation through different assistive tasks.

Most of these works focus on the independent living of the users by providing them with more autonomy. Lee and Riek [27] define autonomy as how much users such as older adults feel in control over their lives. They argue that to support these users’ autonomy, they should have enough control over the robots. Moreover, Chien et al. [14] show that providing elements of personal relevance such as customized designs and settings, as well as positive experiences, can improve the acceptance and use of assistive robots by older adults. Therefore, the customization of the robot by means of preferences may have a great impact on the autonomy that Physically Assistive Robots provide. Preferences may be used to grant more control to the user over how the robot behaves as well as more options for customization and personalization. Other authors have also focused on the need for personalization and adaptation to improve user assistance. For instance, Klee et al. [26]...
provide personalized assistance for hat dressing in a turn-taking manner. To do so, constraints are generated and learned based on human behavior in response to the robot’s repositioning requests. User modeling is used by Gao, Chang, and Demiris [41] to personalize dressing assistance to users with limitations. The user’s movement space is modeled to obtain regions with different reaching capabilities. This space is then used to plan the robot motions when dressing a sleeveless jacket. This approach is basically reactive and the adaptation does not take into account user preferences. Kapusta et al. [24] present a task optimization of robot-assisted dressing, TOORAD, where a plan is generated for both the person and the robot, by using a simulation model including geometry and kinematics of the human, the robot, and the environment. This helps to provide personalized plans for users. A study with six participants with physical disabilities, with the system successfully assisting four of them. The authors state the need for variations on the forms of assistance, which we try to assess in this paper.

Moreover, personalization and user preferences in HRI has also been widely acknowledged in a broader range of topics. Tapus et al. [34] developed a Socially Assistive Robot (SAR) to assist, encourage and interact with post-stroke patients in rehabilitation. The personality of the robot is modified to match the user’s one, modifying elements such as the levels of extroversion and introversion. They also modify robot proxemics and speed and vocal content, similarly to some of our proposed preferences. The authors demonstrate that users prefer the robot that has an adapted behavior to match their personality, leading to improved task performance. Moro et al. [29] propose a behavior personalization method for SAR consisting of the combination of Learning from Demonstration and Reinforcement Learning (RL). In it, the caregivers demonstrate different behaviors for the robot, which are used to learn general behaviors. Then, RL is used to obtain a policy that selects the most appropriate behavior given the user’s cognitive level. Preferences for hand-overs are analyzed by Cakmak et al. [5], where robot and object configurations are adapted to the user preferences that relate to the position and orientation of the object. Lee et al. [28] carried out a long-term field experiment with a snack delivery robot. Participants interacted with the robot over two months and the robot adapted their social interactions. The robot was able to remember the user’s favorite snacks, usage patterns, and robot behaviors. Results showed an increase in rapport with the robot and the users and increased participants’ cooperation. Learning of user preferences based on ratings is performed in [2], where preferences are learned from user ratings over time. Three models of users are defined, which integrate preference profiles. They show its applicability in preferences over light animations in a mobile robot. Chevalier et al. [13] designed a personalized HRI environment for individuals with autism spectrum disorder. The user’s personality (introversion and extroversion) is used to create a model that predicts user preferences by Cruz-May and Tapus [15]. The user evaluation performed shows that the behaviors generated by the model were more preferred by the users. Preferences for social interaction parameters such as distance and speeds are evaluated by Rossi et al. [31], where comfortable stopping distance is evaluated against human pose and user personality. The authors state the importance of learning the preferences in order to adapt the robot behavior, which would foster acceptance by the users and safety feelings. User preferences for robot motion planning are learned by Wilde et al. [39]. User path ranking is used to learn user constraints regarding the task. Similarly, Hayes et al. [21] learn user preferences over continued interaction to improve navigation tasks. Preferences for service robots are studied by Torres et al. [36], where a preference reasoning is proposed. Preferences over service robotics scenarios such as object placement and user tastes are used to interact with the user.

In addition to all this, assistive robots have the problem of acceptance. Society is still reluctant to accept social robots for domestic purposes, as analyzed by de Graaf et al. [16]. The authors state that it is vital to include the opinion of future users in order to adapt the robots to their preferences.
A long-term acceptance study with older adults was performed by Piasek and Wieczorowska-Tobis [30], where high acceptance was demonstrated by users in need of social assistance, with this user segment being open to facilitating technologies. Another study in the same direction was performed by Smarr et al. [33]. In it, twenty-one older adults completed questionnaires regarding their preferences and openness to assistive robotics. The results show a preference for instrumental tasks such as housekeeping or medication reminders but were less open to receiving assistance in tasks such as shaving or hairdressing. A similar analysis is performed by Deutsch et al. [17], where a qualitative study with thirty cognitively-able individuals. Results show many opportunities for home robots, and some user needs that feel threatened by the inclusion of such devices. However, the current generation shift may lead to inconclusive results when experimenting with current older adults. To this end, Gessl et al. [20] study the perceptions and acceptance of future older adults. The study spans 188 users from 20 to 60 years of age, where they found correlations between age, gender, and personality with technology acceptance. The authors state that personality plays a significant role in the acceptance of assistive robotics technologies. Although there may not be a clear link between users’ personality and their preferences, we believe this strengthens the need for personalized and preference-based robotic behaviors, given that robot acceptance may be increased if they can adapt their behavior to match the user needs. Other similar studies, like the one by Biswas et al. [4], try to assess whether older adults are different from young users when interacting with robots, in order to understand how communication preferences change by age. Results show the older people’s preferences in speech over a tablet were more similar to those aged under 21 than the ones in the middle group aged from 22 to 64. Recently, a model to measure Robot Acceptance (the RAM-care model) [37] was proposed based on a survey of healthcare professionals with firsthand experience with the use of robots. The model builds on some technology acceptance models integrating interpersonal values, moral and instrumental ones as notable factors to account for robot acceptance in care settings. It also shows that personal values affect the perceived usefulness of the robots, which would be related to user preferences.

Although many important works have focused on the use of preferences, we are not aware of any of them analyzing whether users can assess the use of their own preferences, which is what we study in this paper. We believe this is a key factor to improve user acceptance and increase the overall interaction and assistance.

3 METHODOLOGY

While it may seem apparent that preferences and adaptation can improve a Physically Assistive Robotics task, there are not many experiments proving so. However, the design of a system able to modify its behavior based on user preferences does not imply such a system will be adaptive and compliant with the preferences as the user would expect, as this is a subjective measure of the user. And, although the user may be able to distinguish the behaviors while comparing them, this would not usually be the desired case for assistive robotics at homes. In that case, the system should be more plug-and-play and managed by non-technical users, without the need of knowing how are the different preferences translated to actual robot behaviors.

Therefore, the main purpose of this study is to determine whether a user can identify when the robot’s behavior is guided by his/her own selection of preferences, given our system and setup. We also want to know if the preferences give a feeling of better assistance. Finally, we want to find out if the behaviors produced by different preferences are observed as different by the users, to see not only whether the user can perceive when their preferences are used, but also when they’re not and the different behavior is observed.

Accordingly, the input parameters will be the chosen preferences by the user. These will be fed into a task planner that will choose the different actions to be executed based on the preferences
The task planner uses methods similar to the ones presented in [8]. With this approach, the planner is free to choose any of the available actions to complete the task, but there is an associated penalization on each action that doesn’t satisfy the user preferences. Therefore, it may be possible that the planner selects actions that do not comply with the preferences in some situations. Our task planning system comprised the ROSPlan framework [10] with the RDDL extension [9] and the PROST planner [25].

The following subsections will detail the hypotheses, setup, scenarios, instruments, and experimental methodology.

### 3.1 Hypotheses

We have devised three hypotheses to assess the impact and usefulness of the use of preferences to modify the robot behavior in assistive tasks.

- **H1**: the preferences modify the behavior of the robot in an expectable way.
- **H2**: the interaction with the robot is more pleasant when preferences are used.
- **H3**: different preferences will produce clearly different robot behaviors identifiable by the user (regardless of their beliefs about the use of their preferences).

H1 will allow us to assess whether the users can determine when their preferences are used. H2 will evaluate if the preferences produce more pleasant behaviors. Finally, H3 is meant to confirm that the preferences do produce different robot behaviors. This is considered independently of the preference’s ability to produce apparent behaviors, which is covered by H1.

### 3.2 Scenarios

We have selected three Physically Assistive Robotics tasks to be carried out by the participants. In what follows, we detail the tasks that have been used to evaluate the use of the preferences along with a description of the preference values that the user could choose for each task. Examples of the setups for the different tasks are shown in Figure 2.

- **Assistive feeding**: The robot feeds the user one spoonful. The spoon is empty for the experiments. The task-related preference is the proximity, which states whether the user wants the robot to be close, far, or at intermediate distance. This is used by the planner to decide whether to use an action that inserts the spoon in the user’s mouth, or another one that lets the user get the food while the robot waits.
- **Shoe fitting**: The robot approaches a shoe to the user, who prepares the foot in front of it, and the shoe is fitted. Then, the robot releases the shoe and the task is finished. The preference to be chosen relates to the movement of the right foot, which could be either nothing, a bit, or a lot/normal movement. Users were allowed to freely simulate an ankle injury that prevented

![Fig. 2. Setup used for the experiments in the three tasks.](image-url)
them to move the foot. The robot behavior changed in the way the shoe was fitted, and in the amount of force that was applied against the foot.

- Jacket (open sleeved garment) dressing: The robot has the garment grasped with two arms, and it dresses the user. Possible behaviors are starting with one arm or the other, or inserting both together. This could be done in each part of the task, be it the sleeve insertion, lower arm dragging, or upper-arm dressing. In this task, the user could choose how much the right arm could be moved, again simulating injury if desired. The possible values were no movement, a bit, or normal movement. This preference guided the behavior of the robot by starting for the right arm or using actions that would not affect the arm movement, either by forcing the user to move or by stretching the arm with the garment.

These tasks are widely addressed by the PAR community, given that they have a potentially high impact on the dependent users’ lives. They represent some of the most basic needs without which a human cannot live with dignity. The selected tasks are still safe enough to prevent user harm when performed by a robot.

Given that we are mainly focused on the behavior changes observed by the users, we have safely implemented the tasks by compromising some elements in favor of safety. Therefore, we have implemented fixed trajectories learned kinesthetically in joint space to prevent any potential misbehavior or unexpected movement of the robot due to sensor faults, misdetections, or potential singularities in the inverse kinematics that could harm the user.

We believe that, although this may hinder a bit the experience of the users, it does not impede them the assessment of the behavior of the robot neither the ability to discern the cases in which the preferences are being used. This does not prevent them to be able to participate in the task and experience the assistance first-hand in a safe environment.

Apart from the task-related preferences, the user could choose two additional preferences for each task (but with the possibility of providing different values per each task). We decided to use the same two preferences to simplify the user’s task. These preferences are:

- Robot speed: defines the speed at which the robot moves, using fuzzy terms. The options were slow, intermediate, or fast. No demonstration of the speeds was performed beforehand, so it was based only on user intuition. The planner selected the speed based on the preferences and the reward function specified in the planning domain. The speed was selected by the planner as a parameter of the movement actions.
- Robot verbosity: specifies whether the robot should speak or not. In our scenarios, speaking relates to informing the user before each task was being performed. An example of speech is: “I am approaching your right hand”. Possible values for the preference are sometimes, always, or never. As this is a preference and not a constraint, the robot may decide to speak even in cases when the preference was not to speak. For instance, it may speak to inform the users that it will not insert the food into their mouth but instead wait for them to get the food while the robot stays still. For the planner, speaking was an extra action that could be executed before a movement action, based on the preferences and the reward function. The action was not needed to complete the task, but the execution of the speaking action when the preferences specified it produced more reward for the planner.

The tasks have been implemented in the robot and behaviors have been designed for each of them. The decision making is performed using task planning, and the preferences are used in the same way as described in [8]. During the tasks, the robot was acting in a fully autonomous way without external interventions.
3.3 Material
The experiments were carried out using two 7 degrees-of-freedom Barrett WAM® robotic arms. The feeding and shoe fitting tasks were using only one arm, while the jacket dressing task employed both of them. The robots prepared to execute each task can be seen in Figure 2. The robot behavior was based on pre-learned trajectories before the experiments began. Each task was split up in different actions (such as “approach to foot” or “insert spoon”), and each action was executed by one or more trajectories (i.e., different approaching movements). The learned trajectories were defined as the standard speed movements, and they were slowed down and speeded up to generate the trajectories for the fast and slow velocities. When executing the tasks, the robot was reproducing the prerecorded trajectories according to the actions chosen by the task planner (action, specific trajectory, and speed). The reward function driving the planner action selection was based on a set of rules that link the actions with the preferences, generating reward when the preferences are fulfilled. As an example, when speed was high, the robot would try to not approach frontally (as we perceived this as being more threatening). For the same reason, this approach strategy was also discouraged (given a negative reward) if the proximity was set to be high. The speed bands were explicitly narrower for the feeding trajectory due to the inherent danger of the task. All the movement actions were executed using compliant controllers that allow for perturbances in the robot movement, as well as stopping on collision.

Additionally, each user was given a single-use spoon for the feeding task. This task also needed a table, a plate, and a chair. No food was supplied to the users, both for health and safety reasons. The shoe-fitting task required a shoe, which was chosen to be a Crocs™-like shoe, and a stool to sit the person in. Finally, the jacket dressing task was using a big size sleeved shirt that would fit all the users and was wide enough to prevent harm by the garment getting stuck while fitting.

3.4 Participants
Thirty participants volunteered to take part in the experiments. The recruitment process was performed through posters and email announcements in departmental mailing lists of the university. The participants ranged from students, researchers, administrative staff, and some relatives of members of the university community. Most of the participants were familiar with technology. All of the participants in the study were healthy, cognitively-aware, and able-bodied adults. Even though those users are not the target population for our applications, we believe the state of the Physically Assistive Robotics systems are not mature enough to be used with potential end-users. We needed some part of flexibility from the user as well as full cognitive capacities to understand the task, the test, and what was their role in the experiment, which was to focus on the robot’s behavior and see whether there were changes and their preferences were used.

Such experimentation in a care home may have been a source of stress for the patients and may have led to inconclusive results if the goal of the task was not understood, apart from possibly generating false expectations to the patients of such care homes. The fact that nowadays we are still far to see real Physically Assistive Robots helping real users made us shift the focus to younger users who may need one of these robots in the future [20]. Thus, we believe this choice of participants does not harm the experiment but helps to understand how potential future users see this kind of systems, and whether the preferences are well understood. Therefore, we specifically targeted the study to healthy people from 20 to 40 years. Those are users who are more familiarized with technology and personalization of devices such as smartphones and, thus, able to understand the goals of the experiment.

3.5 Instruments
We have evaluated the user experience with two questionnaires after the task was performed, and a third one was used to gather the user’s preferences before each task. The presented questions were single-answer multiple-choice questions and 5-point Likert scale questions (ranging from 1 - strongly disagree to 5 - strongly agree).

Figure 3 shows the questionnaire used to obtain user preferences. Before each task, the users were selecting three preferences, one of each type, from the list of available options. For each task,

(a) Preference selection questionnaire presented to the user before each assistive task.

(b) Evaluation questions presented to the user after each assistive task.

Fig. 3. Task-related questions presented to the user before and after the tasks. The users had to mark only one option for each question.
they were selecting a task-specific preference and the speed and verbosity ones. After the tasks were executed, we assessed the perceived use of preferences, pleasantness, and opinion on the potential helpfulness of the task for dependent users. The questions presented to the participants can be seen in Figure 3b.

The users were finally presented with the Almere model questionnaire [22] to assess robot acceptance. Given that the participants were not dependent users, they were asked to answer the questions as if they required assistance to perform the tasks of the experiments. While it is not possible to fully extrapolate acceptance to that of final users, we believe it is worth evaluating the current acceptance. Thus, a study with able-bodied users is a first step towards the validation of the technological approach and allows us to get some insights into the potential future acceptance of these systems. Feelings of rejection towards the robot by able users who can timely react and avoid potentially unpleasant situations would likely indicate that dependent users would not accept the robot either.

3.6 Procedure

The experiments were conducted at IRI’s Perception and Manipulation lab6. The study was designed as a within-subject experiment in which the participants were exposed to different conditions of the same task.

Each user started the experiment by learning about the purpose of the study and the tasks to be performed. The users were informed that the goal of the study was to evaluate if they were able to assess when the robot was using their preferences and when not. They were also informed that the study evaluated the potential use of these kinds of robots and their adaptivity. Regarding the preferences, the participants were told that preferences may appear in the first run of the task, in the second, or in none of them. After learning about the experimental procedure they signed the participant consent form. Then, they were asked about their preferences for the robot behavior, and this was done before each task was executed. All the tasks were performed two times by each participant. In each of them, one execution would use the user’s preferences to guide the robot’s behavior, and the other execution would use explicitly different preferences to the ones selected. There was also a control task for each user in which the chosen preferences were not used, and both executions of the task were using the same set of preferences. This control task should show no perceived differences in the robot’s behavior. The user did not know in which run would the preferences be used and was only told that it may be that there were used in the first execution, the second, all of them, or none. After each task, the user would answer some questions regarding the task, to state where they believed their preferences were used (where options were in the first run, in the second, or dubious). The task-related questions also included some Likert-scale questions assessing their feelings regarding the task and the robot behavior, as stated in Section 3.5.

In order to balance the executions, the following strategy was used. Three possible cases were defined for each task:

- CASE A: The task will be executed two times without preferences.
- CASE B: First execution will use the preferences, the second won’t.
- CASE C: First execution will not use the preferences, the second execution will.

Then, we allocated the cases to the tasks such that each user was experimenting the three cases (one per each task) and balancing between tasks. Therefore, the first user would start with the tasks following cases A, B, and C, and the next one would experiment cases B, C, and A. This way, all the tasks were allocated the different cases for different users. Therefore, it was not possible that one

---

6https://www.iri.upc.edu/research/perception
Are Preferences Useful for Better Assistance?: A Physically Assistive Robotics user study

39:11

Fig. 4. Users participating in the experiment.

The task was always executed without preferences, neither that it included preferences in the same order in all the executions.

At the end of the three tasks, the users completed the survey by answering a randomized set of questions from the Almere model [22].

4 RESULTS

Thirty users (86% self-identified as male) aged between 20 and 39 (mean of 26) years old participated in the experiment. Images of some participating users performing each task can be seen in Figure 4. For a more detailed example refer to the video, where the robot using the different preferences can be seen in action\textsuperscript{7}.

4.1 Preference guessing

The first question the users faced after each task had been executed two times was to assess in which of the trials they believed their preferences had been used. As shown in question (1) from Figure 3b, their options were that the preferences were used in the first trial, in the second, or they were not sure (i.e., no difference found).

\textsuperscript{7}A video with different users and preferences can be seen in http://www.iri.upc.edu/groups/perception/assistivePreferencesEvaluation
The results are shown in Figure 5. As it can be seen, there were 66.67% of correct guesses for the feeding task, 80% for the shoe task, and 70% for the jacket dressing. If joined together, the total amounts to 72.22% of successful guesses in the trial where the preferences were used.

As it can be seen in the plots, most of the cases were successfully identified. The most misguessed case was the one in which no preferences were involved. The observed behavior of the users was that sometimes, even when they were displaying a clear doubt, they were selecting trial 1 or trial 2. Asked after the test, some of them stated that they believed one of the two trials would involve preferences even when they did not notice much difference during the task execution and they were informed in the beginning that it might be that preferences were not present. Even with it, all the cases were correctly identified, being the feeding task the most difficult one. We attribute this to

**Fig. 5.** Results of guessed preference trial per task. The 0 represents doubt/no preferences, 1 represents in the first trial and 2 in the second one.
the fact that being the most dangerous task, the differences in speed and movements, which were less diverse, may have not been that much distinguishable by the users. Anxiety due to the inherent danger of the task may have also made the users less focused on the robot’s actions. To assess significance, we performed a Chi-squared test on the correctly guessed results with a confidence level of 95%, obtaining a p-value of 0.006 < 0.05.

4.2 Pleasantness of the interaction

While our main focus was to determine whether the users can understand when their preferences were being used, we also explored how was the pleasantness perceived in relation to the preferences. Therefore, after each task, the users were also asked about the pleasantness of the interaction when preferences were present. When no preferences were available, the users tended to answer either a low score or middle score to express neutrality. During the experiments and posterior survey, the users were never aware if they had guessed correctly the use of preferences, nor whether their preferences had been used or not. Satisfaction increased regardless of users’ ability to guess correctly.

Figure 6a and Table 1 show the results obtained for the pleasantness of the interaction considering all the answers, while Figure 6b and Table 2 show only the answers in the cases where the preferences were successfully identified (and there were preferences). It can be observed that the pleasantness levels increase when the preferences are correctly identified, as when taking into consideration all the answers there are more low scores, while considering only the guessed results the good scores (values 4 and 5) are more present. The effect is less clear in the feeding task, but there is still a clear increase in the pleasantness. These results would support our hypothesis H2, where the use of user preferences leads to a more pleasant task. Table 3 provides further descriptive statistics on the obtained Likert-scale responses, showing that the results were higher in the cases where preferences were present, also supporting that when preferences are not present the overall pleasantness is lower. A Chi-squared significance test gives us confidence in the results, obtaining a p-value < 0.05 both for the aggregated results of all the tasks and for the shoe fitting and jacket tasks independently.

4.3 Differential behavior

Our last hypothesis, H3, was that the use of preferences would lead to clearly different robot behaviors, independently of the chosen preferences. To assess this, we asked the users about whether the behavior was assessed as different using a 5-point Likert scale question, where 1 means “strongly disagree” and 5 means “strongly agree”. We would expect high scores for cases with

<table>
<thead>
<tr>
<th>FEEDING</th>
<th>SHOE</th>
<th>JACKET</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>20.00%</td>
<td>23.33%</td>
<td><strong>36.67%</strong></td>
</tr>
<tr>
<td>4</td>
<td><strong>43.33%</strong></td>
<td><strong>33.33%</strong></td>
<td><strong>36.67%</strong></td>
</tr>
<tr>
<td>3</td>
<td>20.00%</td>
<td>30.00%</td>
<td>26.67%</td>
</tr>
<tr>
<td>2</td>
<td>6.67%</td>
<td>10.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>10.00%</td>
<td>3.33%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 1. Frequencies of the 5-point Likert scale for the pleasantness of interaction. Considering all results.

<table>
<thead>
<tr>
<th>FEEDING</th>
<th>SHOE</th>
<th>JACKET</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td><strong>40.00%</strong></td>
<td><strong>43.75%</strong></td>
<td><strong>56.25%</strong></td>
</tr>
<tr>
<td>4</td>
<td>20.00%</td>
<td>25.00%</td>
<td>37.50%</td>
</tr>
<tr>
<td>3</td>
<td>26.67%</td>
<td>18.75%</td>
<td>6.25%</td>
</tr>
<tr>
<td>2</td>
<td>6.67%</td>
<td>12.50%</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>6.67%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 2. Frequencies of the 5-point Likert scale for the pleasantness of interaction. Only correctly guessed state when preferences were present are considered.
preferences, while the cases in which preferences were not present should show disagreement. Figure 7 and Tables 4–5 show a comparison of the obtained results when preferences are not present and when they are applied. As it can be seen, there’s much more confusion (a score of 3) and disagreement when preferences were not used, while the use of preferences shows more agreement in the difference of behaviors between the first and second trial. Similarly, Table 6 shows the same trend, where cases using preferences were identified as having clearly different behavior. In this question, there is no difference regarding the correct guessing of the use of the preferences. A Chi-squared test for this question also showed significance for the shoe-fitting, jacket dressing, and aggregated results, for which we can consider our hypothesis accepted.

4.4 Potential usefulness of Assistive Robotics

The results in this section are prospective. First, our targeted users did not require this kind of assistance. However, as most people can relate or think of cases in which PAR may be of use, we asked them about their beliefs on the potential uses of these assistive robotics tasks. Second, our experiment used a robot prototype far from being market-ready and, as explained before, some real-time motion adaptation was removed to favor safety.

The collected answers are summarized in Figure 8. Both charts show a clear opinion in favor of the use of Physically Assistive Robots to help dependent people. Most of the users agreed that the use of preferences such as the ones proposed in this paper has the potential to improve the

<table>
<thead>
<tr>
<th></th>
<th>FEEDING</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OP</td>
<td>NP</td>
<td>SGP</td>
<td>OP</td>
<td>NP</td>
</tr>
<tr>
<td>MEAN</td>
<td>3.85</td>
<td>3</td>
<td>3.8</td>
<td>3.8</td>
<td>3.3</td>
</tr>
<tr>
<td>MEDIAN</td>
<td>4</td>
<td>3.5</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>MODE</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics for the pleasantness. OP are the results only considering the cases where preferences were present. NP are the cases where preferences were not present. SGP are the results when there were preferences and the users correctly guessed in which trial.

![Fig. 6. Results of the pleasantness experienced by the users. Observe that 4 and 5 are the dominant answers overall. The results improve with an increase of agreement answers when only the successfully guessed state is taken into account, meaning that the use of preferences increases the task pleasantness.](image-url)
Are Preferences Useful for Better Assistance?: A Physically Assistive Robotics user study

(a) Behavior difference when preferences are not present

(b) Behavior difference when preferences are present (independent of correct guess)

Fig. 7. Results of behavior difference per task with and without preferences. Note that when preferences are present, the users agree on an observed behavior difference, while when no preferences are present the users get more confused.

<table>
<thead>
<tr>
<th>FEEDING</th>
<th>SHOE</th>
<th>JACKET</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>20.00%</td>
<td>0.00%</td>
<td>20.00%</td>
</tr>
<tr>
<td>4</td>
<td>20.00%</td>
<td>10.00%</td>
<td>20.00%</td>
</tr>
<tr>
<td>3</td>
<td>20.00%</td>
<td><strong>50.00%</strong></td>
<td><strong>30.00%</strong></td>
</tr>
<tr>
<td>2</td>
<td>10.00%</td>
<td>20.00%</td>
<td>10.00%</td>
</tr>
<tr>
<td>1</td>
<td><strong>30.00%</strong></td>
<td>20.00%</td>
<td>20.00%</td>
</tr>
</tbody>
</table>

Table 4. Frequencies of the 5-point Likert scale for the difference of behavior. Only when no preferences were applied

<table>
<thead>
<tr>
<th>FEEDING</th>
<th>SHOE</th>
<th>JACKET</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>15.00%</td>
<td>30.00%</td>
<td>30.00%</td>
</tr>
<tr>
<td>4</td>
<td><strong>45.00%</strong></td>
<td><strong>40.00%</strong></td>
<td><strong>50.00%</strong></td>
</tr>
<tr>
<td>3</td>
<td>20.00%</td>
<td>25.00%</td>
<td>20.00%</td>
</tr>
<tr>
<td>2</td>
<td>20.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>0.00%</td>
<td>5.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 5. Frequencies of the 5-point Likert scale for the difference of behavior. Only when preferences were present regardless of user’s guess.

<table>
<thead>
<tr>
<th>FEEDING</th>
<th>SHOE</th>
<th>JACKET</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>OP</td>
<td>NP</td>
<td>SGP</td>
</tr>
<tr>
<td></td>
<td>3.55</td>
<td>2.9</td>
<td><strong>3.6</strong></td>
</tr>
<tr>
<td>MEDIAN</td>
<td>4</td>
<td>3</td>
<td><strong>4</strong></td>
</tr>
<tr>
<td>MODE</td>
<td>4</td>
<td>1</td>
<td><strong>4</strong></td>
</tr>
</tbody>
</table>

Table 6. Descriptive statistics for the behavior difference. OP are the results only considering the cases where preferences were present. NP are the cases where preferences were not present. SGP are the results when there were preferences and the users correctly guessed in which trial.

assistance for those in need. In a more general question, they also confirmed that those applications could be helpful for dependent users based on their opinion. As it can be seen in Figure 8a, the majority of the users agreed that the preferences may improve the assistance. Similar results are observed in the helpfulness, where users also confirmed that these tasks can provide effective assistance to users in need.

Finally, we have used the Almere model [22] to get further insights into the acceptance of this kind of robots. The questions were provided randomized and adapted to the experiment at hand,
appealing to the users’ imagination in some parts, where they were told to put themselves in the role of a dependent user. The Social Presence (SP) construct was not included in the questionnaire given that there was no conversation between the robot and the user.

The obtained results are summarized in Table 7. Even though not all the constructs had results above the significance level, we believe they give us good insights into what the users think about the system. Above all, we want to notice the constructs as Anxiety (ANX - evoking anxious reactions) which shows the lowest scores. The Attitude Towards Technology (ATT) shows positive feelings about the applicability of the technology, being the one with the highest score. Trust and Intention To Use (ITU) show a wider range of opinions, while the Perceived Usefulness (PU) and Perceived Adaptivity (PAD) give insights of good acceptance of the system and perception of a useful and adaptive system, which is even enjoyable (PENJ).

5 CONCLUSIONS

Physically Assistive Robots may have a huge impact on future society and home care for dependent people. In this paper, we have carried out a user evaluation with thirty healthy subjects in order to assess the effect of the use of preferences to modify the behavior of the robot. Our experiment intended to determine whether the users were able to assess when a robot is using their chosen preferences in an assistive task and when it was not. Moreover, we evaluated the impact of such preferences in the perceived pleasantness of the task, and also the opinions on future usability of the proposed assistive tasks. Nonetheless, we understand having able-bodied participants may

<table>
<thead>
<tr>
<th>Construct</th>
<th>Mean ± SD</th>
<th>Alpha</th>
<th>Construct</th>
<th>Mean ± SD</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANX</td>
<td>2.71 ± 0.85</td>
<td>0.72</td>
<td>PEOU</td>
<td>3.81 ± 0.56</td>
<td>0.46</td>
</tr>
<tr>
<td>ATT</td>
<td>4.33 ± 0.54</td>
<td>0.68</td>
<td>PS</td>
<td>3.02 ± 0.74</td>
<td>0.64</td>
</tr>
<tr>
<td>FC</td>
<td>3.51 ± 0.74</td>
<td>0.31</td>
<td>PU</td>
<td>3.99 ± 0.59</td>
<td>0.55</td>
</tr>
<tr>
<td>ITU</td>
<td>3.02 ± 1.08</td>
<td>0.91</td>
<td>SI</td>
<td>3.87 ± 0.64</td>
<td>0</td>
</tr>
<tr>
<td>PAD</td>
<td>4.08 ± 0.44</td>
<td>0.20</td>
<td>Trust</td>
<td>3.42 ± 1.05</td>
<td>0.78</td>
</tr>
<tr>
<td>PENJ</td>
<td>3.92 ± 0.70</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Results of the Almere model [22] analysis.

(a) Perceived improvement of the assistance of dependent people by the use of preferences

(b) Perceived helpfulness of the applications for dependent people

Fig. 8. Answers for the improvement of the preferences and the perceived helpfulness of the assistive tasks.
be a potential limitation of the study when looking into the care setting, but it provides valuable insights on the use of preferences in assistive robotics scenarios.

The obtained results allow us to confirm that most users are able to determine when their preferences are being used, meaning preferences are self-explanatory even when the users did not have any previous experience to compare with. We have also observed that in the absence of preferences, some users may try to assess that there were preferences even when there were not. And even in that case, the results show a clear understanding by the user of when the robot was using the preferences and how those were manifested. When it comes to pleasantness, we can conclude that the use of preferences led to a more pleasant sensation for the users. Furthermore, the users show clear agreement in that applications such as feeding, shoe fitting, and dressing assistance may be helpful for dependent users, and preferences and personalization would improve this assistance. Finally, regarding the acceptance, we can’t have conclusive results given that our system was a prototype and our users were not end-users, but our preliminary results on the topic show that the users can trust the system without having much anxiety over it. The system was intuitive and easy to use, for which most of the users reported that they would use a system like this in case of need.

Given the obtained results, we can confirm our hypotheses and conclude that our use of preferences for Physically Assistive tasks modifies the behavior of the robot in an expectable and legible manner and that doing so results in a better experience for the user. We have also observed that different preferences produce clearly different behaviors, as expressed by the users. We have gotten important insights on how preferences can play a role in physical assistance. However, we believe this study should not end up here, and a follow-up would be helpful to get further intuition by extending it to users with some kind of disability or dependency degree, but still with full cognitive capacities in order to understand the goals of the study. This will also help to better assess whether anxiety conditions of real users could affect their acceptability for the robot, for which we provided preliminary results. For this, a more complete system would be needed, with full vision capacities to interact and adapt to the user and safe control strategies to prevent any harm.

ACKNOWLEDGMENTS

The authors would like to thank the users who participated in the study. This work has been approved (report number 025/2019) by the Ethical Committee of the Spanish Council of Scientific Research (CSIC). We thank the committee for their evaluation and comments regarding this experimental evaluation. This work has been supported by the ERC project Clothilde (ERC-2016-ADG-741930), the HuMoUR project (Spanish Ministry of Science and Innovation TIN2017-90086-R) and by the Spanish State Research Agency through the María de Maeztu Seal of Excellence to IRI (MDM-2016-0656). Gerard Canal has also been supported by the Spanish Ministry of Education, Culture and Sport by the FPU15/00504 doctoral grant and the CHIST-ERA project COHERENT (EPSRC EP/V062506/1).

REFERENCES


