Exploring Interaction Modalities and Task Allocation for Household Robotic Arms

Sascha Herr & Tom Gross
Human-Computer Interaction Group
University of Bamberg, Germany
<firstname>.<lastname>@uni-bamberg.de

Michael Gradmann & Dominik Henrich
Chair for Applied Computer Science III
University of Bayreuth, Germany
<firstname>.<lastname>@uni-bayreuth.de

CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892456

Abstract
Human-robot interaction in private households demands easy and natural interfaces, as well as an understanding of areas of application. Two fundamental issues are multimodal interaction (i.e., multifarious ways of communication with the robot) and task allocation (i.e., division of labour between the user and the robot). We report on an explorative interview study that gathered users' requirements for multimodal interaction and gained understanding for task allocation in private households.

Author Keywords
Human-Robot Interaction; User Study; Requirements; Multimodality; Task Allocation; Private Households.

ACM Classification Keywords
I.2.9 Robotics: Commercial robots and applications.

Introduction
In 2007, Bill Gates predicted that robots would soon follow the same development as computers, in that they would become common in every home [6]. So far, there has been limited success for domestic robots; predominantly in the field of vacuum cleaning robots [10]. To pave the way for a proliferation of domestic robots, a reasonable direction is to enhance their utility

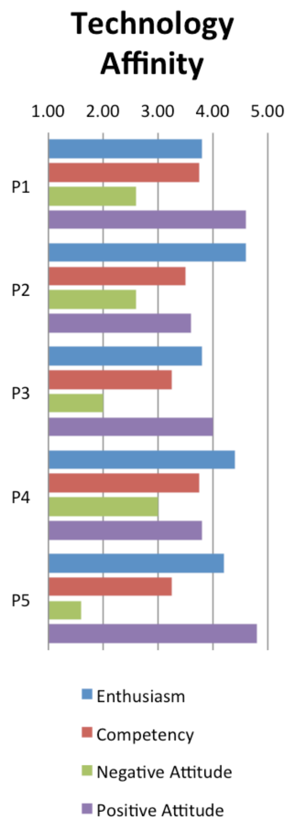

Figure 1. Participants' technological affinity.

by offering more versatile models such as a robotic arm that is not restricted to dedicated activities, but can autonomously accomplish and cooperatively support a broad spectrum of tasks.

This approach combined with the individuality of household contexts poses challenges to developers in human-robot interaction (HRI). First, a clear understanding of the areas of application is necessary to develop appropriate functionalities. Quantitative approaches to identify personal user expectations towards robots provided long lists of expected activities [2, 3], but the diverse nature of household tasks and the importance of the task context require a more in-depth investigation that also takes task allocation into account. Specifically, the potential role of the robot might differ for tasks [1, 4], i.e. users might want to delegate some tasks completely, while they might expect assistance in others.

Second, as household features drastically differ from home to home and users are likely to hold diverse preferences regarding the execution of tasks a natural and easy interface to control and program a robot is required [2]. One way to achieve this is the use of multimodal interfaces. Multimodality describes multifarious ways of communication (e.g., via gestures, gazes, or speech), which users are familiar with. Although large-scale quantitative studies investigated general expectations regarding types of interaction modalities [9], the diverse contexts with respect to household configuration and specific user situations demand a closer explorative investigation.

This paper explores these challenges with a qualitative approach. Our contribution is twofold: First, we derive requirements for several specific interaction modalities for a robotic arm. Second, we investigate task allocation and gather an understanding for user motivations and task characteristics in household HRI.

Method

We conducted semi-structured interviews to gather in-depth information from prospective users of HRI in private households.

Participants

We used diversity and convenience sampling to recruit five participants (2 female, 3 male). The sampling criteria were: own household; clear end-user perspective (no technical background); and interest in technology from a consumer's perspective. We employed a standard questionnaire with 5-point Likert items measuring the affinity to technology (TA-EG [7]). Participants exhibited similar competency with technology, but showed differences with regard to their attitudes and enthusiasm. Strikingly, participants in their 40s (P1, P3) were not as enthusiastic as participants in their 30s (P2, P4, P5). P2 and P4 exhibited the least difference with regard to positive and negative attitudes towards technology, while P5 scored extreme values with both the lowest negative attitude score and highest positive attitude score (cf. Table 1 and Figure 1 for an overview of sample demographics and participants' technology affinity).

Material

Our interview guide had instructions for the interviewer as well as questions in a suggested order. The 20 questions followed interview guidelines, posing open questions and employing prompts and probes when appropriate [8]. We collected ideas and opinions

Part.	Gender	Age	Occupation
P1	F	44	Office Administrator
P2	M	31	Contract Manager
P3	M	43	Teacher
P4	M	30	Pricing Manager
P5	F	31	Sleep Lab Assistant

Table 1. Participants.

Figure 2. Example picture for introducing participants to the topic.

towards interaction in general and modalities specifically. Regarding task allocation, we focused on kitchen work, deskwork, and handicraft—typical areas where a desk-mounted robotic arm can help. For each of these contexts we asked participants how they would allocate tasks between them and the robot.

Procedure

Predominantly the interviews took place in the participants' homes (P5's interview took place in our laboratory). They lasted 90-120 minutes and were video recorded for later transcription. The interviewer continuously took notes after obtaining informed consent. To provide a reference frame and introduce the topic to the participants we provided them with three pictures and a video where a lightweight robotic arm (KUKA LBR; cf. Figure 2) performed household activities. An arm was chosen to avoid unexpected social interaction intentions common with creature-like robots [5]. Following the face-to-face interviews, we transcribed them and conducted a thematic analysis [8], i.e. we identified interesting sections and marked them, extracted meaningful units as codes, and clustered these in order to find major themes.

Results

Through our analysis we extracted requirements, expectations, and opinions on interaction modalities and task allocation, which we summarise and illustrate with participants' statements below.

Interaction Modalities

We asked openly, but also probed common modalities such as force guiding, demonstration, gestures, gazes, and speech. We focused on modalities that allow communication by means of the user's body.

Force guiding involves direct manipulation of robotic limbs to record movement patterns. Participants conceived this as intuitive. However, there was some scepticism regarding the complexity to teach simple tasks due to the user's natural understanding of pressure sensitivity that robots are expected to lack. *P2: "I sometimes don't know myself, how much pressure I need to open a bottle of water."*

Demonstration involves showing movements that users expect the robot to imitate. It was seen as a very good option to teach complex movement patterns. *P4: "It would be the easiest to have something like a glove and I just do something and the robot memorizes this. I could imagine that. At least for complex things I would find that more comfortable than [force] guiding it."* It was also mentioned that the robot could learn from mere observation of the user. *P2: "I think that further development will lead the robot to observe you and determine on its own what to do."*

Gestures involve users' arm movements that are visible and signal commands to the robot. Participants had the impression that gestures were viable but might be cumbersome. *P5: "When I point at the pen on the floor and show it a grasp gesture… I think before I would do that, I would have picked up the pen myself 3 times faster."* Privacy concerns due to the constant video capture necessary for gesture recognition were voiced as well. *P2: "If I have to communicate with gestures with the robot, it means that I am surveilled all the time. I think I would feel uneasy about that."*

Gazes involve eye movements that the robot recognizes and interprets. Participants doubted that recognition of their gaze for object reference is accurate. *P2: "Well, now I look at you [the interviewer], but I actually focus on the bottle behind you. That would be difficult for the robot to recognize."* Furthermore, they argued that voluntary control of the eyes is difficult on the user's part. *P4: "I think gazes are the one thing you have the least control over."*

Speech involves commanding the robot by voice. Participants assumed that a translation from human terms to robot understanding might be complex and that cognition of object and location references must be built in. *P2: "…the robot has to have a high degree of understanding already built-in. […] I cannot just tell him 'go to this position of the table, or that position of the table', since I don't exactly know that myself."*

Task Allocation

Based on our analysis we conceptualise task allocation as three categories from a user perspective (compare [1]): complete delegation, collective work, and solitary work.

Complete delegation of tasks can be performed without the users' direct interaction and must be programmed beforehand. Tasks that participants want to delegate were characterized as time-intensive, cumbersome, unpleasant, and after-products of main activities. Participants often mentioned tasks that fall into the category of general cleaning and tidying up. Participants also brought up garbage-related as well as maintenance tasks (e.g., closing drawers on a file cabinet and filling up the rinse aid for the dishwasher). *P4: "Cleaning the drain is an activity that I could do without. It is not fun and it is cumbersome. […] This is just something I have to do since I use my kitchen extensively."*

Collective work is characterised by a direct and synchronous interaction between user and robot. Tasks that users want to perform with the robot's assistance involve unpleasant movements (e.g., plugging in cables under the desk), safety issues (e.g., working with a circular saw), or situations where the robot is not trusted to handle it on its own (e.g., filing). Additionally, users mentioned activities where a third hand is beneficial, such as handing over objects when the user is occupied (e.g., tools or dishes) and providing support (e.g., holding wooden planks while the user saws them or holding presents tightly for application of scotch tape). *P2: "It could even hold the lamp up there while you search for the screwdriver."*

Solitary work denotes areas where users are sensitive and reserved regarding the involvement of a robot. Users want to do these tasks without a robot. The types of tasks are user-specific: Three of our participants said there should be no involvement of a robot in food preparation—as it is a valued activity that establishes connection to nourishment and where appreciation can be shown through effort. *P2: "Cooking is an activity that you have to do yourself to have some relation to the food."* In contrast, two participants would delegate cooking completely. *P3: "Many people love to cook, but we don't. For our children, we want healthy food on the table, so we have to, but we just do not like to cook."* Furthermore, participants voiced that humans are inherently superior to robots with certain tasks (e.g., cooking, handicraft, dishwashing). *P4: "I am convinced that my chaotic dishwasher loading system is superior to any other system."* Participants also expressed reservations with tasks that involve private data (e.g., filing), tasks that are joyful (e.g., crafting), relaxing (e.g., knitting) or result in pride (e.g., finishing initially

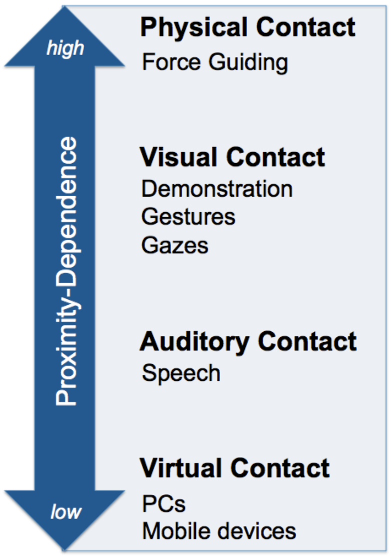

Figure 3. Proximity-dependence of interaction modalities.

unpleasant activities such as filing or physical work). *P4: "I find it to be an extremely liberating experience to do the filing. You have an incredibly full desk and at the end of the year you get to file everything in orderly fashion. Then you have everything neatly in folders and you are the happiest person."*

Participants mentioned several instances where they would change their allocation if the situation required it. Participants would call on robot assistance for otherwise solitary work tasks if special occasions arose such as cooking larger quantities. *P2: "Potato salad for 20 people? Peeling potatoes is not fun anymore!"* One participant voiced that if she were not satisfied with the robot performance, she would change delegated tasks to solitary work. *P1: "When I observe that the robot destroys my wine glasses, I would have to revise my stance and would not let the robot do that anymore."* Also worries about being bored with too much delegation came up. *P1: "When I have the feeling that the robot does everything and I don't do anything, I would say: Boring! I have to do something again!"*

Discussion

The analysis of transcribed interviews suggests various implications for the design of future robots for private households, with respect to user attitudes, proximity-dependence, combination and interplay between modalities and task allocation.

User Attitudes towards Household Tasks

In our interviews we received the strong impression that the participants' attitudes towards specific household tasks determined their task allocation, as well as their willingness to spend time on teaching the robot. The participants' personal values and beliefs seem to influence how they decided what the robot was allowed to do. This also relates to the assessment of technology affinity. Our most opinionated participants (P2 and P4) are also quite 'balanced' in their attitudes towards technology, in that they exhibit the least difference between positive and negative attitudes towards technology (cf. Figure 1). Furthermore, the notion of human superiority to robots emerged when participants doubted the technology to work to their satisfaction or individual tastes. A one-size-fits-all approach therefore cannot be employed; users need to easily configure the robot for their specific needs. The willingness to invest time for teaching the robot highly varied (simple tasks: 1min–4h, complex tasks: 2min–weekend). Users may have no motivation to teach the robot if the time requirement for doing the task on their own is short. In general, participants would invest more time for tasks that they want to delegate completely compared to collective work tasks. The attitude towards household tasks seemed to play a major role: the extent of personal dislike for a task is the main factor for willingness to invest time in teaching.

Proximity-Dependence of Modalities

Participants argued repeatedly that the interaction modalities imply different proximities to the robot (cf. Figure 3). Physical contact requires direct contact with the robot (e.g., force guiding requires the user has to manipulate robotic limbs directly). Visual contact requires the user to be within the visible area of the robot. Demonstration, gestures, and gazes fall into this category. As no direct physical contact is needed, they offer a lower proximity-dependence but still require the user to be in relatively close proximity to the robot. Auditory contact requires the user to be within a distance where acoustic signals are received. Speech falls into this category. Virtual contact, with which the robot is controlled via mobile devices or computers offers the least dependence. Proximity-dependence should be kept in mind when designing multimodal interaction, as it also encompasses perceptual reach and different privacy-related biases of users.

Modality Combination
Combinations of modalities are reasonable for different interaction objectives. Speech is suitable for identifying objects with auditory signals when combined with other modalities. Force guiding in conjunction with speech can convey the names of objects as the robot is guided to grab them. Gaze recognition combined with speech can link objects on which visual attention is focused with spoken names. Speech can also serve as an attention-grabbing signal that puts the robot in a wake state for gesture interaction. Gaze recognition can provide preparatory information that reduces the space of available reference objects and also supports object distinction when gestures are used. Furthermore, cross-teaching new modality usage (e.g., saying "move cup to table" while simultaneously force guiding the action) could render teaching more fun and efficient.

Modality and Task Allocation
Some combinations of modality and task allocation are more suitable than others. In complete delegation no synchronous interaction happens. Actions have to be programmed beforehand, for instance, via force guiding for complex movements. In collective work tasks, users may be occupied with their hands and not be able to force guide the robot or use gestures, so that the support for speech and gaze is recommended. For solitary work tasks that involve privacy-related material, the user should be able to put the robot in a sleep state by gesture or speech that makes the robot arm less obtrusive and visually present by having it align perpendicularly to the wall or hide behind blinds.

Conclusions
Our study revealed diverse findings. The priority of input modes depends on factors such as characteristics of the task at hand and the user attitudes. Users have strong opinions on allocating tasks to robots and feel a need to be able to dynamically alter allocations. Overall the realisation of user needs depends on circumstances, most importantly on the perceptual reach between users and robot.

Future household robots need to be both pre-equipped with fast possibilities of recognizing and interacting in their environment with users, and be easily personalisable to respond to user-specific needs.

The limited sample size and cultural diversity prohibit from generalising our results. However, the explorative nature of our study revealed potentially interesting research opportunities regarding task allocation, user attitudes, modality constraints and chances in household applications for robotic arms. Larger scale quantitative studies could investigate these topics more specifically and establish generalisability. Further research is also needed to address the issue of time investment for teaching more thoroughly, as well as the robustness of our findings for other robot types.

Acknowledgements
We thank the members of the Cooperative Media Lab in Bamberg, and the participants of the study. Part of the work has been funded by the Technology Alliance of Upper Franconia (TAO).

References

1. Klaus Bengler, Markus Zimmermann, Dino Bortot, Martin Kienle and Daniel Damböck. Interaction Principles for Cooperative Human-Machine Systems. *it-Information Technology Methoden und innovative Anwendungen der Informatik und Informationstechnik*, 54, 4 (2012), 157-164.
2. Maya Cakmak and Leila Takayama. Towards a Comprehensive Chore List for Domestic Robots. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction - HRI 2013 (3-6 Mar*, Tokyo, JPN). IEEE, 2013.
3. S. N. Copleston and G. Bugmann. Personal Robot User Expectations. University of Plymouth, Plymouth, UK, 2008.
4. Kerstin Dautenhahn, Sarah Woods, Christina Kaouri, Michael L. Walters, Kheng Lee Koay and Iain Werry. What is a Robot Companion - Friend, Assistant or Butler? In *Proceedings of the International Conference on Intelligent Robots and Systems - IROS 2005* (2-6 Aug, Edmonton, Alberta, CA). IEEE, 2005, pp. 1192 - 1197.
5. Agostino De Santis, Bruno Siciliano, Alessandro De Luca and Antonio Bicchi An Atlas of Physical Human–Robot Interaction. *Mechanism and Machine Theory*, 43 (2008), 253-270.
6. Bill Gates A Robot in Every Home. *Scientific American*, 296 (2007), 58-65.
7. Katja Karrer, Charlotte Glaser, Caroline Clemens and Carmen Bruder Technikaffinität erfassen - der Fragebogen TA-EG [Measure affinity to technology: The questionnaire TA-EG]. In A. Lichtenstein, C. Stößel and C. Clemens (eds.) *Der Mensch im Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme (ZMMS Spektrum)* 22, (2009), 196-201.
8. Nigel King and Christine Horrocks *Interviews in Qualitative Research*. Sage, Thousands Oaks, CA, 2010.
9. Céline Ray, Francesco Mondada and Roland Siegwart. What Do People Expect from Robots? In *Proceedings of the IEEE/RSJ 2008 International Conference on Intelligent Robots and Systems - IROS 2008* (22-26 Sep, Nice, France). IEEE, 2008.
10. Florian Vaussard, Julia Fink, Valérie Bauwens, Philippe Rétornaz, D Hamel, Pierre Dillenbourg and Francesco Mondada Lessons Learned from Robotic Vacuum Cleaners Entering in the Home Ecosystem. *Robotics and Autonomous Systems*, 62, 3 (2014), 376-391.