Effects of Robot Capability on User Acceptance

Elizabeth Cha, Anca D. Dragan and Siddhartha S. Srinivasa

The Robotics Institute, Carnegie Mellon University. {lizcha,adragan,siddh}@cmu.edu

Index Terms—**Acceptance; capability; human-robot interaction; manipulate; perception; personal robot**

I. INTRODUCTION

Although personal robots hold great promise, they face many barriers to being accepted and adopted in our homes [1]–[3]. As these robots are expected to perform personal tasks, their burden of trustworthiness is far greater than that of service or industrial robots.

Several studies have explored our acceptance of robots, studying factors such as natural interaction via speech [4], gaze [5], and appearance [4], expectation setting via apologizing [6], and predictability of behavior [1].

However, all of these studies ground the *capability* of the robot by either showing participants images and videos of one specific robot, or not showing them anything at all, leaving it to their imagination. A key limitation of these approaches is that they do not inform robot designers of the capability that they should strive for in their robots.

Of course, the obvious answer is that surely better capability would produce greater acceptance. But, the details are important. How capable should a robot be so that you would trust it to bring you coffee? Would it be different from the capability it would need to clean up your room?

Furthermore, we discovered that the *perception* of capability could be as important as the true capability itself. For example, one of our participants was worried that a faster, more capable, robot might spill coffee because it was moving too quickly.

In this paper, we explore how capability affects what users want out of a personal robot. We focus on two questions: 1) How does a robot's capability affect user's acceptance of a home robot? and 2) What types of tasks in the home would users want assistance from a robot of a certain capability?

We explored these questions by conducting a pilot study where we manipulated *task factors*, many of which were inspired from Beer et al. [2]. However, in addition, we *deliberately* manipulated the *robot's capability*, for example, we speed up or slowed down videos, and added extra seemingly redundant motions.

Our results indicate that, on average, the more capable robot was more acceptable than a less capable one, with subjects citing speed of motion and the lack of pauses as contributing factors. While this was not surprising, the small effect size was. Some subjects complained that the robot was still not fast enough, while others trusted the robot but still did not want it to perform their task because they would rather do it themselves. Finally, one user *disliked* the speed, complaining that the robot somehow seemed less deliberate and methodical.
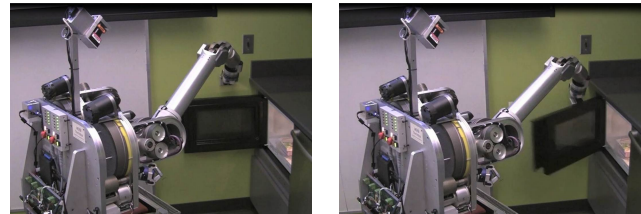
(a) Robot 1: slow, guarded moves (b) Robot 2: fast, direct moves

Fig. 1. We manipulate robot capability.

Our study reveals that manipulating capability is useful, but still very challenging, and opens the door for further research on robot capability.

II. EXPERIMENTAL DESIGN

We designed a pilot survey study to explore the effect of robot capability on the acceptance of a personal robot helping with various tasks. The subjects watched videos of HERB (Fig.1) and answered questions based on these videos.

Manipulated Variables: We had 8 sub-studies, each manipulating the *robot's capability* and a *task factor*.

We manipulated the robot's capability by creating two separate videos, both showing HERB performing the same task (microwaving a meal) in different ways. The first video (*Robot 1*) portrayed HERB's current capabilities: HERB moved at its usual speed and performed its usual series of extraneous motions ("guarded moves"), to help localize its hand in the environment e.g. before closing the door. HERB took almost 5 minutes to complete the task. The second video (*Robot 2*) portrayed a more capable HERB, moving much faster and skipping the guarded moves (e.g. pressing the button directly). This version of HERB took less than a minute for the task.

Each of the 8 sub-studies manipulated a task-related factor using tasks we devised or inspired from [2]:

1) Collaboration: robot-only tasks vs. collaborative tasks with the user (clean up the table vs. clean up the table with me)
2) Proximity: tasks at a distance vs. in the personal space of the user (retrieve my jacket vs. help me put it on)
3) Specific Instruction: tasks performed in the way specified by the user vs. in the robot's own way (put the flour back where I say vs. where it finds appropriate)
4) Personal Preferences: generic tasks vs. tasks with strong user preferences (clean up the table vs. organize the pantry)
5) Damage-Prone Items: tasks involving breakable, spillable objects vs. robust tasks (bring mug with coffee vs. refreshment from refrigerator)
6) Important Items: tasks involving generic objects vs. important ones (bring generic mug vs. favorite mug)
7) Important Living Entities: similar to above, but with living things (caring for plant vs. caring for pet)
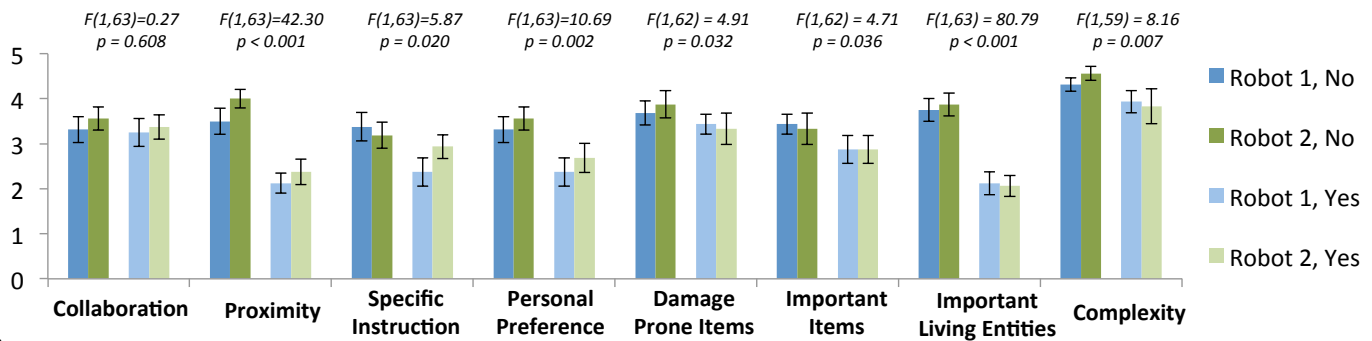
Fig. 2. ANOVA analysis for each of the 8 sub-studies. In the legend, Yes and No refer to the task factor (e.g. collaborative and non-collaborative task.

8) *Complexity*: easy task vs. hard task (placing clothes in hamper vs. doing laundry)

We also devised a few additional tasks that are relevant to our work on HERB, e.g. meal preparation or book organizing. **Dependent Measures**: After watching the videos, subjects were presented with a statement of the form: "Based on the capabilities of Robot X, I would want the robot to do Y ." We collected each subject's *rating* of the statement on a 5-point Likert scale from strongly disagree to strongly agree.

Subject Allocation: We chose a within design, where all subjects provided ratings on all tasks for both robots. We surveyed 16 subjects (10 male, 6 female) recruited through Mechanical Turk, aged 22 to 46 ($M = 30.06$ years, $SD = 7.45$). No subject reported high familiarity with robots. All subjects were from the Unites States and had over 95% approval ratings. Additionally, we asked the subjects a few control questions to ensure they had watched the videos carefully.

Hypotheses:

1: *Robot capability* has a significant positive effect on *rating*.
2: Each *task factor* has a significant effect on *rating*.

III. RESULTS

Overall across the data, the more capable robot had a significantly higher rating ($M = 3.21, SD = 1.31$) than the less capable ($M = 2.98, SD = 1.29$): $t(312) = 3.77, p < 0.001$.

For each sub-study, we performed a repeated-measures factorial ANOVA with the task and the capability factors. Fig.2 summarizes the results from this analysis: all but the first task factor were significant, and the robot factor was not significant (we discuss in the next section possible causes).

The mean differences are in the expected directions: users did not trust the robot with tasks in their proximity, would prefer to specify the task details, would not trust the robot with their important items, etc. Users wanted the robot to care for their child the least, and wanted the robot to do laundry- and trash-related chores the most.

IV. DISCUSSION

On average, robot capability swayed the users from slightly below accepting the robot (2.98) to slightly above it (3.21). Most subjects did prefer the more capable robot, citing its efficiency and precision: They said Robot 2 "seemed to know what to do without a lot of pausing", "was able to handle tasks

in a quick, efficient manner", "was faster than robot 1 so I would be more likely to let it help me with chores", etc. One subject said "I chose answers that I previously did not want Robot 1 to do for me," citing speed as the reason. Another commented that "I would be more comfortable having a robot like this doing these tasks for me." Users were also willing to pay significantly more for Robot 2 ($p = 0.001$).

For some subjects, Robot 2 was "still not fast enough" for their needs. This indicates that acceptance by some users will require performing tasks at close-to-human speeds.

One subject disambiguated between trusting the robot and wanting it to do a task: "I would rather do my own laundry." This provides a useful design improvement for our study: different questions to evaluate different nuances of acceptance.

Speed is not enough to manipulate robot capability. First, a faster robot is sometimes perceived as less capable. For one user, Robot 2 "seemed less deliberate and methodical."

Second, capability goes far beyond speed. A lot of users deemed the faster robot still incapable of many tasks. They said that they "still wouldn't want it doing things that require delicate care" and that they would be afraid of the robot making a mess or breaking their favorite mug. Users also remark that robots lack situational awareness, they "don't understand whats going on around them", they "can't think or reason", making them unsuitable for taking care of pets or loved ones.

Overall, our pilot took a step towards answering what amount of robot capability is enough for various tasks. It revealed, however, that manipulating robot capability, albeit important, is a challenging experimental design problem. We look forward to discussing possible approaches for improving this manipulation with the HRI community.

REFERENCES

[1] K. Dautenhahn *et al.*, "What is a robot companion-friend, assistant or butler?" in *IROS*, 2005.
[2] J. Beer *et al.*, "The domesticated robot: design guidelines for assisting older adults to age in place," in *HRI*. IEEE, 2012, pp. 335–342.
[3] J. Sung *et al.*, "Sketching the future: Assessing user needs for domestic robots," in *RO-MAN*, 2009.
[4] J. Goetz *et al.*, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in *RO-MAN*, 2003.
[5] C. Muhl and Y. Nagai, "Does disturbance discourage people from communicating with a robot?" in *RO-MAN*. IEEE, 2007.
[6] M. Lee *et al.*, "Gracefully mitigating breakdowns in robotic services," in *HRI*. IEEE, 2010, pp. 203–210.