AN ABSTRACT OF THE THESIS OF

Dilruba Showkat for the degree of Master of Science in Computer Science

presented on June 6, 2017.

Title: Gender Differences in Robot Teleoperation

Abstract approved: ____

Cindy Grimm

Gender issues have recently received increased attention in human robot interaction (HRI). Because robots are becoming part of our homes and daily lives, it is important to understand how different groups of people use them. To the best of our knowledge, almost no research has been done that investigates gender differences in users information need, information processing strategy, self-efficacy, tinkering and their impact in human robot interaction. This thesis investigates these four aspects by examining object manipulation task from gender perspective using a humanoid robot (PR2). We used both qualitative and quantitative approaches for cross validation and methodological triangulation. Our experimental results show that females asked for more information before using the robot than males *($p = 0.0002$)*. Females processed information comprehensively and males processed information selectively *($p < 0.001$)* for using the robot. Males showed greater self-efficacy than females *($p = 0.0002$)*. Males tinkered more with the robot

than females $(p = 0.0021)$. We found that tinkering was positively correlated $(p = 0.0068)$ with task success and negatively correlated $(p = 0.0032)$ with task completion time. Tinkering perhaps led to males greater task success and lower task completion time with the robot. Findings from this research can be useful for making design decisions for robots and open new research directions.

Gender Differences in Robot Teleoperation

by

Dilruba Showkat

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 6, 2017
Commencement June 2018

Master of Science thesis of Dilruba Showkat presented on June 6, 2017.

APPROVED:

Major Professor, representing Computer Science

Head of the School of Electrical Engineering and Computer Science

---

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Dilruba Showkat, Author

ACKNOWLEDGEMENTS

First of all, thanks to almighty for blessing me with love, grace and kindness. I am highly thankful to Dr. Cindy Grimm who supported me throughout every step of my masters by guiding my research, meeting deadlines and coordinating with others. I would like to thank Dr. Margaret Burnett for her continuous support throughout my research from an early stage of study design to execution. Without her help and guidance this research was almost impossible. I would like to thank Dr. Anita Sarma for helping me learn the best from her courses and class projects. The study materials helped me develop and improve my background knowledge I needed to complete my research successfully. I would like to thank Matthew Rueben and Austin Allen Whitesell for helping me with the PR2. Finally, I would like to thank my family for supporting me in my career endeavor.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF ALGORITHMS

LIST OF APPENDIX FIGURES

LIST OF APPENDIX TABLES

Chapter 1: Introduction

In the next few years, robots are expected to become a significant part of many households, helping elderly senior citizens for self care, socialization, child care as robotic assistant and companion [25]. As this happens, it becomes apparent that robots will help improving our lifestyle. Therefore, it is becoming more important for researchers to investigate on how different groups of people perceive and use robots as robots are becoming an essential part of our daily life as well as participating in social situations. Human perception of robot is important because it tells us what is expected from robots (how the robot should behave? what makes the robot trustworthy? etc.). Therefore, the social-psychological processes in human robot interaction (HRI) requires further investigation.

Previously, factors such as how a robot should approach a person sitting on a chair, the influence of a robots voice, robots gender, the robots facial features and how all these plays out for different groups of people (males and females) has been investigated. Existing research show gender differences in the robot's social presence, social facilitation, disclosure and persuasiveness [77, 79]. These researches studied how a robot can persuade different individuals to disclose their private information, how different individuals perform in math task, how receptive people were to the robot suggestions [29, 70]. Furthermore these research findings are significant because they helped us understand how a robot is perceived in different situation from the perspective of different groups of users (males and females). It is important that we know about users expectations because it is the user who are going to use it.

There is a large body of research that has been done in gender differences. Gender differences has been studied in communication, education, creativity, human computer interaction (HCI), human robot interaction (HRI), web psychology and many more [1, 21, 51, 68]. There is numerous applications of gender in providing and improving users experience with a product or service. For example, Researchers developed gender recognition algorithm to improve the development of real-world assistive technology [92], for example to notify an inattentive driver (who could be male/female). Identifying gender difference has numerous application, for example a robot interacting with a human (appropriately addressing Mr. or Mrs.) [64]. Furthermore, researchers proposed a systematic approach called GenderMag (persona) to evaluate usability of problem-solving software for gender inclusiveness issues [18]. GenderMag persona's helped companies to discover and fix gender issues in their software to make it more accessible [43].

In this research, we are taking our first step towards identifying gender differences in robot teleoperation (manipulating robot from a distance) by investigating users information need, information processing strategy, self-efficacy and tinkering. Though we already know robots are perceived differently by males and females but still yet we don't know whether there exists any gender differences in the aforementioned factors. Furthermore, almost no existing research addressed this problem. Outcome from this research can help us design robots that will bridge gaps if any discrepancy found.

In this study, we are employing empirical research based investigation to understand whether there exists any gender differences while manipulating a robot for the first time. This research is important because robots will become part of our daily life working as a helper and assistant, we would want to know as a designer whether our design of robot is accessible to everyone or not so that we can provide the right information at the right time. The outcome from this research will help us design and develop robots that improves accessibility, enable greater user experience and will also open new research direction.

Our study was designed and structured with the following research questions in mind.

RQ1: Are there any gender differences in the information need before using a robot that users have never used before?

RQ2: Are there any differences in the information processing style across gender when learning how to use a robot?

RQ3: Are there any gender differences in the self-efficacy that impact efficient use of the robot?

RQ4: Are there any gender differences about tinkering with the robot?

Chapter 2: Related Work

Robots are already in our homes for exp. Amazon's Alexa, and they will continue to grow at their fast paced acceptance rate in various applications such as in surgery, heath care, engineering education and more [16, 22, 63, 91]. These diverse applications demands more experimentation on how these technologies has been received across different demographics (age, gender, education, culture etc.)

Gender issues has recently become a buzzword both in human robot interaction (HRI) and human computer interaction (HCI). In this chapter, we will give background details of existing researches related to our study. Differences between HCI and HRI has been investigated by [37, 78]. The study indicates that HRI concerns with systems that use complex and dynamic control system, and operate in highly changeable environment. First of all, we will introduce socio-technical aspect of human-robot interaction and the influence of gender in HRI. Secondly, we will briefly describe future applications of robots. In third, we will describe existing researches related to gender differences in various domains.

2.1 Socio-technical aspect and influence of gender in HRI

In the introduction, we discussed socio-technical aspects involving human and robot. By changing robots appearance such as gender (voice, facial features etc.) have shown to have significant influence on how the recommendation provided by the robot was received by male and female, how the robot is perceived by males and females, and how the robot impacts task performance (easy and hard math). The study reported that males tend to think of robots as more human like whereas females think of robots as more machine like. As a result, males reported to feel socially facilitated by the robot while performing arithmetic tasks, whereas females did not [66, 70, 77, 79].

Furthermore, research show that by changing robots persona we can gather different level of information, for example a study reported that males expressed more information to the female robot and females expressed more information to the male robot. In another research, males and females participants reported that they find opposite sex robot to be more trustworthy and engaging. Researchers also found gender differences in the negative attitude toward robots [66, 70, 77, 79]. All these findings are interesting from the perspective of latest voice enabled technologies such as Amazon's Alexa, Google's Assistant etc. These bots are female voice enabled (Siri for iPhone, Microsoft's Cortana) as the basis of conveying information [87, 81]. Furthermore, previous research confirms that using only vocal cues within a machine is enough to bring sex based stereotype responses even though the environment and circumstances in which the robot operates could be different [23, 78]. When robots are given some human like attributes, people can easily relate to them. Even though we consider robots/bots as machines we still tend to use he/she when addressing them [81]. Therefore, all these researches and products (bots) that we are using on a daily basis are a direct evidence that gender differences needs to be studied more to design and implement products that makes us happy.

2.2 Future application of robots

In this section, we are going to give brief description of possible future applications of robots.

2.2.1 Robots managing households

Household robots consist of cleaning robots, elderly citizen care robots, nursing robots, entertainment purpose robots and so on. Research indicate that using robot therapy increased social interaction among the elderly residents in a care house [86]. Our future homes will be surrounded by these robots. These robots face installation challenges because our homes are unstructured and humans are generally unpredictable in nature. This makes it hard for the cleaning robots to autonomously navigate around. Consequently making it difficult to decide when and where to or not to clean. Neato botvac connected robot vacuum and iRobot Roomba 980 has been considered the best vacuum robot in 2017 [71]. Neato botvac is wifi equipped and app enabled on IOS or Android devices. With all these latest technology comes the challenge of security and privacy vulnerabilities [30, 49, 86].

2.2.2 Robots in education and teaching

Another important application of robots is in the education. Researchers found that robots can have great applicability in educational technology by studying the role, type and behavior of the robots [63]. By investigating social supportive behavior for a robot (iCat) tutoring students to learn a language (artificial language "Toki Pona") showed increased learning efficiency [75]. While computer based education showed proven benefits as the emergence of fast growing internet service, but it cannot provide mentor tutor relationship. On the contrary, a robot can actually provide more engaging and interactive user experience. Latest inventions such as Dash, mBot allows children to learn programming (using Google Blockly), Arduino and robots is yet another huge development in education [49].

2.2.3 Robots in assisting information sharing

This is an interesting category of robots that are used for dedicated task such as helping users with directions, providing instructions, serve as museum guides etc. These robots can be used for greetings in offices as well to provide useful information that are less changeable. The problem in the lab environment with controlled lab testing is that it can fail to mimic real unstructured interaction with humans but for regular common interaction robots can serve humans [49, 76].

2.2.4 Robots in communication

As robots are sent in critical places such in wars, rescue, hospitals, what becomes really important to understand is how humans and robot communicate with each other [49, 76, 85]. Robots are also useful in remote communication. This kind of remote communication has been conducted between/among humans via video-conferencing techniques which highly depend of facial expression, body language and other non-verbal cues. Though is solves many problem of long distance communication but it still cannot solve all problems such as side conversation and pointing object is not possible. These problems can be eliminated using robots with enhanced intelligent features.

There are many other future opportunities and robot applications as mentioned in earlier sections such as in scientific exploration, for search and rescue, health care and in elder homes. In all these instances robots will be socially interacting with humans. The social robots are useful in different social context and are distinguished in four categories: 1) socially evocative 2) socially receptive 3) social interface and 4) sociable [14]. Understanding these categories are important because it tells us which robot to use in any particular application.

As we are hoping to build a bright future with numerous applications of robots, it is inevitable to consider ethical issue that comes with it [55, 56]. There is a risk and fear among workers that machines will replace humans in the workplace. For example, with the emergence of autonomous vehicles or self driving cars will replace drivers, another example would be how uber, lyft impacted taxi driving industry [27]. Advancement in robot application will not only impact job sector but also safety [34]. A research studied 107 users attitude towards self-driving cars reported that individuals were concerned with liability, cost of the car, and loosing control of the car. The study showed that males were more concerned with liability and less concerned with controls compared to females. Cost was a concern for both groups of users. Researchers also emphasized that there is a need for further development to improve aging drivers experience [13, 44, 90]. We have to make sure that the price we pay for improving our lifestyle must not be more than the price we pay fixing the problems caused by developments. Therefore, we need to enhance our understanding of how different groups (age, gender, education, ethnicity, race etc.) of people are impacted by these developments.

2.3 Gender differences in various domains

In this section, we will describe several facets of gender differences that has been studied widely in social psychology, education, technology and in many other domains [12, 18, 35, 54, 57].

2.3.1 Information processing style and gender

Males and females use different information processing strategies according to selectivity hypothesis [20]. Males style of attention to any task is by discrete segments of configuration, while females style of attention to the whole configuration. Males tend to engage in selective/heuristic based processing making use of single cues that are highly available and most noticeable or important in the current context to make a single inference. On the other hand, females tend to engage themselves in gathering all the available cues as the basis of judgment for information processing. Therefore, females approach to information processing is comprehensive, effortful and complete analysis of the situation. This can be explained by the structure of activities males and females engage in as well. The activities females engage in exhibit high structure which requires group feedback, individual instructions. Whereas males engage in activities having low structure requiring task initiation and leadership [28, 39, 40, 48, 61, 62].

2.3.2 Self-efficacy and gender

Social cognitive theory [5] suggest that self-efficacy or an individual's personal judgment about their own capabilities plays a crucial role in the choices that they make, the amount of effort they put, and task retention when faced with adversity. Later researchers report that males showed to have more self-efficacy, less math anxiety and higher performance score compared to females for mathematical problem solving [60, 67]. Researchers studied self-efficacy in excel and report that females had lower self-efficacy than males about their abilities to debug spreadsheets and were reluctant to accept new software features. Though using these features may be useful for task success but they choose not to try those software features [7, 8]

2.3.3 Tinkering/playfulness and gender

Cognitive playfulness or tinkering demonstrated high test performance in a field study of full time employees [58], students also benefit in their scientific understanding as they tinker or play with tools [46, 52]. Males are likely to tinker more than females was found in education literature [84]. By investigating end-users debugging spreadsheet researchers report that males tinker more than females, but females benefit more as they tinker whereas males tinkering was not indicative of effectiveness [9, 17]. Considering the benefits of learning by tinkering, several projects such as Arduino, Raspberry Pi provides tools for hardware and software tinkering [11, 73].

2.3.4 Attitudes towards risk and gender

Research confirmed by a meta analysis of 150 studies that there is a greater risk taking attitude among males compared to females at different level of age and wide variety of task [19, 36, 42]. Later researchers proposed a scale that can be used to measure risk taking in financial decisions, health, recreational, ethical, and social decisions. By testing the proposed scale with undergraduate students researchers found that females were more risk averse than males [24, 88]. But this did not follow in case of social decisions. This scale is useful for the hiring process in startup companies.

2.3.5 Motivations and gender

In education literature researchers report that students who are intrinsically motivated (takes on tasks for the enjoyment or the learning itself) rather than extrinsically motivated (takes on task for usefulness or some external reward such as grades) tends to do better [12, 54, 57]. Existing literature also show that males tend learn technology for the enjoyment of itself, whereas females are motivated by what they can achieve with it [17, 60].

2.3.6 Personality and gender

Males and females also differ in their personality traits. Alan et. al showed by meta analysis that males are more assertive, having high self-esteem compared to females. Females showed higher anxiety, tender-mindedness with respect to males. These findings in personality traits were constant across different demographics (age, educational level etc.). Some researcher believed these differences were related to our biology [33, 35].

Chapter 3: Experimental Design and Methodology

For this research study, our experiment was designed with research questions in mind such that the data collected allows for both qualitative and quantitative data analysis. This is a between group (male and female) study. Our experiment was carried out in three phases or sessions. We will describe each of these phases in the following sections. The three phases were:

1. Pre-task session

2. Task session

3. Post-task session

When participant enters the study room, we asked them to read through the informed consent and give verbal consent. Participants can disagree if they will. The researcher was present throughout all three sessions. Let's reflect on our study procedure, then we will explain all three steps in detail.

3.1 Study Procedure

Entire work flow of our study is given in Figure 3.1

1. Participant entered the study room and researcher briefly introduced them to the robot (Figure 3.2).

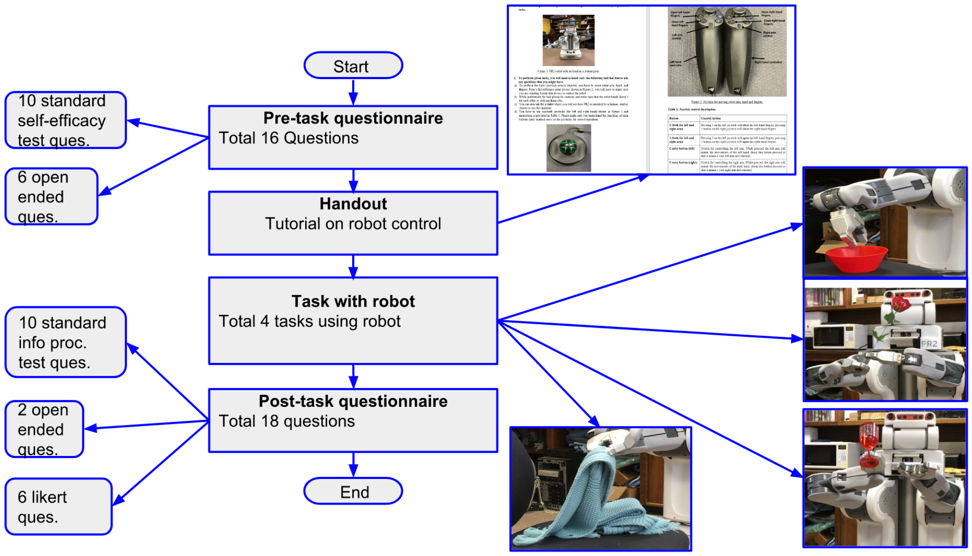

Figure 3.1: Work flow of our research study.

2. Participant reads the consent form and gave verbal consent.

3. The participant filled out the pre-task questionnaire.

4. Task session

   (a) Researcher provided the handout containing tutorial on how to control the robot (Appendix C).

   (b) Researcher provide the joystick controllers (shown in Figure 3.4).

   (c) Researcher placed the object on the chair.

   (d) When the participant was ready, researcher explained the task associated with the object and constraints or any assumptions that were made for the task. Task description and sequence is provided in Table 3.2.

(e) The participant stand behind the reference point (shown in Figure 3.3). The participant stands face to face with the robot at a distant of 4 ft.

(f) Participant completed the task (under all constraints) associated with the object.

(g) Repeat steps (c-f) until all objects (Figure 3.5) are covered.

5. Data recording was completed.

6. Participants were asked to fill out post-task questionnaire.

7. Go to step 1 for the next participant.

3.2 Participants

Participants were recruited via flyer distribution. Flyer was posted by the researcher on and off campus to cover wider range of population. We did not restrict out study participation by any age group or gender. But only students from Oregon state university showed interest in our research study. Communication and scheduling for the study was done through email. Participants were given $15 for their time and commitment. Overall 12 participants (6 males and 6 females) took part in the study and 5 participants in the pilot study. Most of them were undergraduate and masters students from engineering and social science department at Oregon State University. Participants demographics can be found in Table 3.1.

Participant label	*Gender*	*Age*	*Degree major*	*Video gaming experience (in years)*
P01	Male	19	Computer science	10
P02	Female	24	Public health	0
P03	Male	20	Computer science	10
P04	Female	21	Computer science	1
P05	Female	21	Art	10
P06	Male	20	Computer science	11
P07	Female	21	Psychology	0
P08	Male	20	Computer science	8
P09	Male	19	Computer science	5
P10	Male	19	Computer science	13
P11	Female	24	Computer science	0
P12	Female	18	Nuclear engineering	3

Table 3.1: Participant's general demographics.

3.3 Robot Used in the Experiment

We used Personal Robot (PR2) from willow garage [38] for our study. It is shown in Figure 3.2. It is a humanoid robot platform for research and development in robotics. It is durable and does not require any hardware and software implementation from scratch so we can focus on new innovation and application.

3.4 Experiment Setup

At first, the researcher starts the PR2 to make sure it is up and running and ready for use in the experiment. This process of initial setup sometimes took about an hour (after a complete reboot) or more. Participants were asked to operate the robot by standing behind the reference point. Participants were farcing the robot

Figure 3.2: Personal robot (PR2) used in the study (in action).

and that is why they had to mentally adjust the mirrored hand movements, i.e. robots left hand mimicked his/her left hand but as in mirrored image. This has somewhat made the tasks more complicated. The study took place in the personal robotics lab at Oregon State University. During the experiment the lab was quiet, and only researcher and participant were present near the experiment area.

3.5 Pilot Study

We carried out pilot study with 5 participants to identify study design problems. As suggested by [65], 5 users should be sufficient to identify most of the problems within the study design. After the pilot run we had to update our list of objects used in the experiment, tasks and task sequence. During the pilot run the tasks were very simple and showing users how to use the robot control made the tasks even easier. So we decided to have our participants learn on how to use the robot my themselves using a handout. To better understand user behavior we finally come up with a mix of simple and complex tasks. The tasks were chosen such that they were hard yet doable. We also made changes to the questions asked in the pre and post task sessions. Pilots study with small number of users helped us find problems in our study design and thus served a major role in the eventual success of our research.

3.6 Pre-task session

In this session, we asked participants to fill out a questionnaire. The questionnaire started by asking participants demographic information: gender, video gaming experience, age, degree program or major. These information were asked for later statistical analysis. This questionnaire consisted of 16 questions in total. Complete questionnaire can be found in Appendix A.

The first 10 questions were asked to measure participants self-efficacy for using the robot. We used the standard self-efficacy test questionnaire proposed by Compeau and Higgins [26] after slight modification. These modifications were done such that the questionnaire fits the task specific to the robot. We did not change the scaling. The 10 questions were 10 point likert-scale type questions. Response can be anything between (1-10) where 1 indicates "Not at all confident", 5 indicates "Moderately confident", and 10 indicates "Totally confident". A snippet of the questionnaire is given below:

"I could complete the task using robot ..."

Q-1 "... if I had never use a robot like this before" 1 2 3 4 5 6 7 8 9 10

Q-2 "... if I had only the robot manual for reference" 1 2 3 4 5 6 7 8 9 10

The rest of the 6 questions were open ended questions targeted to find answers to several other usability factors. We asked one question about users information need for using the robot, next question asking for their choice between handout and video, one question asking how they will use those information (which they respond to previous question), one question about their motivation of using the robot, one question asking tinkering with a hardware, and the last question about risk taking attitude towards the robot. These questions were asked before the task session to understand participants thought process.

3.7 Task session

In the task session, participants used the PR2 robot shown in Figure 3.2 to manipulate several objects. They were instructed to think-aloud during the task. Think aloud protocol is well known for understanding and elicitation of users mental models for solving problems [45].

3.7.1 Tutorial

Participants were provided with handout and video tutorial. Both of them convey the same information. The purpose of this tutorial was to help participants to get familiar with various robot controls. Even though researcher was present during the entire task session, researcher did not help the participants directly when they asked ”what's that” type of question. This was done on purpose to better understand users information processing style and other factors that we were studying in our research. Researcher helped participants when robot hands were hitting each other. Researcher made sure the safety of the robot and participants. Participants were controlling the robot from a distance such that the participant was free from any danger caused by the robot. Moreover, our tasks did not require the robot to move.

1. ***Handout:*** The handout (Appendix C) contains pictorial representation of PR2 and joystick controls available to the participants to complete given task. The participant stand behind the reference point, reference point is shown in Figure 3.3. Joystick controls shown in Figure 3.4 demonstrated feature description of the controller on how to open/close fingers and moving the left/right arm controls.

2. ***Video:*** The video contained complete example on how to move an object (a fluffy toy). It showed how a participant moves the robot arm using joystick control, failed once and then successfully picked up the toy.

There was no demonstration or tutorial on how to grasp an object from the experimenter side. This was useful to understand users information processing as they use a robot they manipulate a robot they have never used before. The entire session was audio and video recorded. If a participant was stuck (if the robot hand is stuck or s/he is nervous) during the task, they can ask the researcher for their query.

3.7.2 Task Description

Task session took no more than 20 - 25 minutes on an average. Overall, it took 30-40 minutes for a complete session. After pilot study we decided to use four objects having elasticity property, so that they could stretch. It allowed the

Figure 3.3: Reference point.

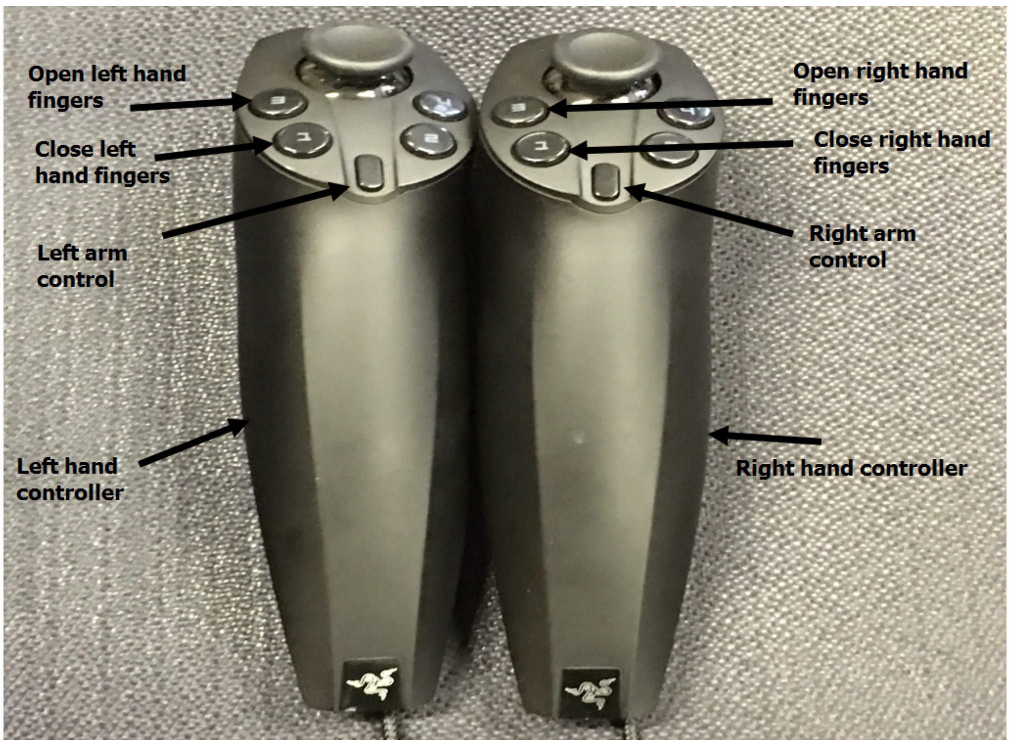

Figure 3.4: Joystick controllers for using PR2.

Figure 3.5: Objects used in the experiment.

robot hands to be safe during high torque when an object is pulled by both hands. Objects used in the experiment is shown in Figure 3.4. Objects and task associated with each object, assumptions are provided in Table 3.2. Furthermore, we added assumptions/constraints with each task (3rd column in Table 3.2) to make sure that our participants had a realistic experience. Because in real life for example users will try their best not to drop a glass, in other words will try best to succeed in each task instead of failing. These assumptions helped us mimic real world experience. Allowing both simple and complex task would help us determining participants information processing strategy better.

Object Name	*Task Description*	*Assumptions*
Bowl	Empty bowl. Picking it up and put it down	Its made of glass
Rose	Pick it up by touching stem only and pass to other hand	It is real flower
Wine glass	Its half filled with drink. Have the robot drink from it	Made of glass
Scarf	Folding it nicely	Its his/her favorite

Table 3.2: Object and task description.

3.7.3 Why object grasping task using robot?

Humans are efficient in grasping objects, whereas robots are not. Even in the perfect experimental setting one in four, i.e. 25% of grasps fail [3]. To enable robust robot grasping, leveraging human grasping adaptation techniques to robots appears to be appealing, because human grasping is optimal and accurate [2]. But there are several limitations to the process of transferring human grasping techniques to robots, such as: 1) lack of time efficient data capture techniques 2) due to limited number of user participation in the study which directly leads to lack of generalizable result and 3) most of the grasping heuristics does not generalize across different humans, objects or environment setting [32]. That's why in this research, we decided grasping task to be solved using PR2 robot to understand self-efficacy, information processing strategy and willingness to tinker.

3.8 Post-task session

The post task questionnaire (Appendix B) contains 18 questions in total. First 10 question were 7-point likert scale type of questions. This is a standard questionnaire where the first five questions tests for comprehensive information processing style and the rest of the five questions tests for systematic information processing [80]. We modified the questionnaire slightly such that it fits the robot interaction task. We did not change the scaling. Response can be anything on a 7-scale between (1-7) where where 1 indicates “Slightly”, 4 indicates “Moderately”, and 7 indicates “Absolutely”.

The questionnaire is then followed by 2 open ended questions asking about what information they used and how did they use those information.

The rest of the 6 questions were 10-point likert scale questions. We asked them 3 questions to measure their after task self-efficacy (1 indicates “Not at all confident” and 10 indicates “Totally confident”). One self task performance evaluation question (1 indicates “very bad” and 10 indicates “very good”). One question about risk behavior (1 indicates “Not at all” and 10 indicates “very much”). Last question is about how did they enjoy the task (1 indicates “Not at all” and 10 indicates “very much”).

Chapter 4: Data Analysis and Results

In this chapter, we are going describe how we analyzed our experimental data, and the methodologies we used in detail. 12 participants took part in our study. We were open in our study across gender, and this was done by asking participants to write down their "gender" instead of asking whether they were male or female. Participants identified themselves as males and females. We had 6 males and 6 females. Participants were on average 20.5 years old (males = 19.5 and females = 21.5), SD = 4.73. Statistical analysis of our background data on video gaming experience showed significant gender differences (Mann Whitney W = 3, p-value = 0.0185). We qualitatively code education of our participants as follows: Arts = 1, Engineering = 2. Statistical analysis revealed that there was no significant gender differences between genders in education (Mann Whitney W = 9, p-value = 0.0705). In most of our test experiment, we used video gaming experience as covariate in ANCOVA analysis and gender (male and female) as factor/categorical variable. We rank transformed our data before applying ANOVA, this is similar to non-parametric ANOVA. We also used Fisher's exact test. Our results are coded by a single experimenter. What makes our results trustworthy is that we validated our results with methodological triangulation and non-parametric tests. Non-parametric test and methodological triangulation strengthens and validates our research outcome. We will discuss and answer each of our research questions

(RQ's) in the following sections.

4.1 Information need and gender in robot interaction (RQ1)

RQ1: ***Are there any gender differences in the information need before using a robot that users have never used before?***

Null Hypothesis (H_0): *There is no gender differences in the information need before using a robot.*

In this RQ, we wanted to know what information users want to know to before using a robot a to grasp and manipulate several objects that they had never used before. For simplicity, we considered three types of information needs. They are 1) strategy, 2) feature and 3) functionality from previous research study [50].

To answer this RQ, we asked participants an open ended question in the pre-task session questionnaire (Appendix A). We asked the following question with a context:

"Imagine that you are asked ***to move a delicate expensive flower vase*** on the table using the PR2 robot. *What information do you need to complete the task using robot?*."

We will describe coding and data analysis methods in the following subsections.

4.1.1 Information need coding strategy

We used qualitative coding to code each responses in three aforementioned categories. Here we will discuss our coding scheme (how and why) with examples.

1. ***Strategy:*** Whenever a response explicitly talked about what would be a suitable process or how to carry out a task we recorded this as a strategy element. For example, consider the following response questions: "How to turn it on?", "How to control it?", "How do I move the arm? How do I open and close the claw? How do I rotate the claw?". All these instances were coded as strategy because these responses directly asked about how to do a particular task. 83.33% of all participants reported information need for strategy.

2. ***Feature:*** When a statement explicitly expressed about what a specific feature (keys, color, markings, buttons) does, we coded this response in feature category. For example consider the responses "What keys to use?","Which button do each task?" is coded as feature because it explicitly queried about keys and buttons. 6 out of 12 i.e. 50% of all participants asked for feature related information.

3. ***Functionality:*** Statement which explicitly query about specific function of a robot hand/arm/control part (what it does). For example responses as "How fast it rotates", "How sensitive the controls are?" is recorded as functional information because they directly talks about specific function of

Code	*Description*	*Examples*
Strategy	Asks about a preferable procedure	“How it works?”
Feature	Talks about a particular feature	“What it uses to grab stuff?”
Functionality	Asks about robot arm/wrist etc.	“How does its hands function?”

Table 4.1: Information need before using PR2 robot coding strategy (response from open-ended question.

the robot and robot parts. 66.67% of all participants asked for functionality related information.

We did not find any responses that could be place in outside of these three categories. See the summary of information need in Table 4.1.

4.1.2 Information need and gender: by the number of categories of questions asked

Using criteria mentioned in Table 4.1, we count information need (strategy, feature, functionality) for each participant as they occur in the responses (open-ended question discussed in section 4.1) and code the count for each participant. Frequency of information need is summarized in Table 4.2. All responses can be be found in Appendix D (Table D.4, D.5, D.6).

We performed ANOVA on our rank transformed data. One-way ANOVA (F (1,10) = 32.472, p-value = 0.00019874 ***) revealed significant gender difference in information need (by number of categories of question asked). Based on all the results, we reject our null hypothesis H_0. And we accept alternative hypothesis that there

Gender	*Strategy*	*Feature*	*Functionality*	*Total*
Males	4 (66.67%)	1 (16.67%)	2 (33.33%)	7 (29.16%)
Females	6 (100%)	5 (83.33%)	6 (100%)	17 (70.83%)
Total	10 (83.33%)	6 (50%)	8 (66.67%)	24 (100%)

Table 4.2: Information need summary before using the PR2 robot (responses from open-ended question). Each entries in the table indicates number of people who have asked for each categories of information need.

is significant gender difference about the information need before using a robot (by number of categories of questions asked). From table 4.2 we can see that of all the responses, 70.83% information were asked by females, and only 29.16% information were asked by males. Even in this small population the gender gap was evident about the information requirement before manipulating a robot (shown in Figure 4.1).

4.1.3 Information need and gender: by the number of questions asked

For this experiment, we count all responses (the same open ended question discussed in section 4.1) in each category (strategy, feature and functionality) and performed ANOVA on our rank transformed data. ANOVA revealed significant gender difference ANOVA (F (1,10) = 10.796, p-value = 0.0082093 **) in information need (by number of question asked). This result is shown in Figure 4.2.

Thus, we reject the null hypothesis and accept that there exists significant gender difference in information need (by number of questions asked) before using the

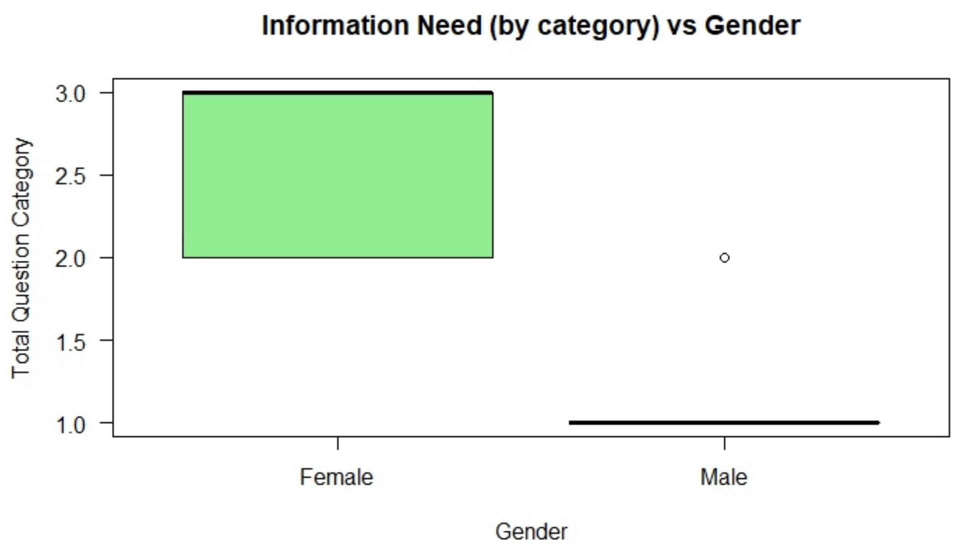

Figure 4.1: Gender difference in information need (by categories of questions asked) before using PR2 robot.

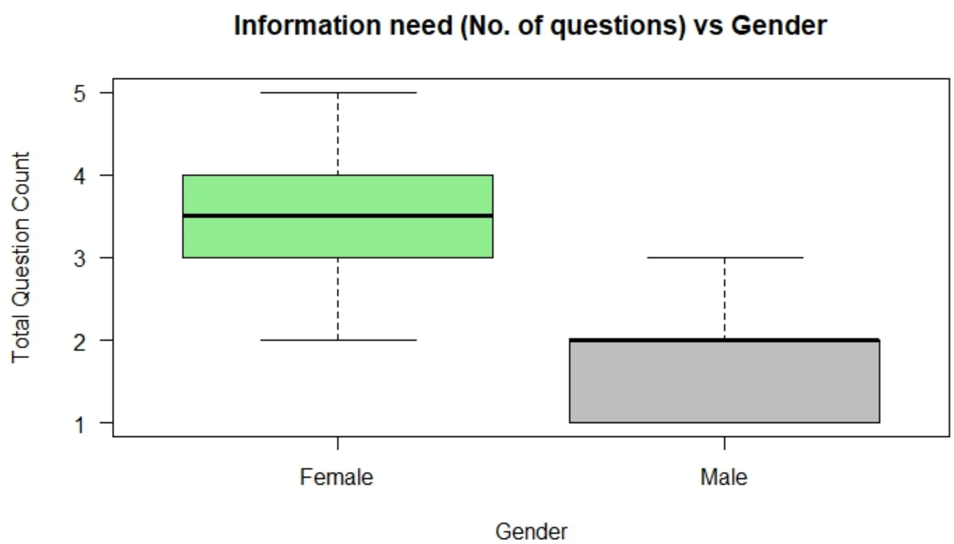

Figure 4.2: Gender difference in information need (by the number of question asked) for using PR2 robot.

robot.

In summary, there exists significant gender difference about the information need both by categories of question asked and the number of questions asked before using the PR2 robot. Our results indicates that female participants asked for more information to make their decision. On the contrary, male participants as heuristic based processor asked for information that made more sense in the current context to achieve their goal.

Post-task session:

We asked the following open-ended question in the post-task session (Appendix B) to identify missing or surprise information need. We asked them to answer based on the final task. *"What information did you use to complete the task?"*. Responses include robot feature, function and strategy related information, for example "I remember that 1 was to the position the fingers and 3 was to release them". We found no new information need (information that they used but users did not mention before) so we did not include these responses in our data analysis.

4.2 Information processing style and gender in robot interaction (RQ2)

RQ2: ***Are there any differences in the information processing style across gender when learning how to use a robot?***

Information Processing Style	*Male*	*Female*
Comprehensive (Sum)	61	111
Comprehensive (Average)	10.16	18.5
Selective (Sum)	125	28
Selective (Average)	20.83	4.66

Table 4.3: Information processing style result summary when learning how to use PR2 robot. Sum and average of responses from information processing style questionnaire.

Null Hypothesis (H_0): *There is no gender difference in the information processing style while learning how to use a robot.*

4.2.1 Measuring information processing style from post task-session

From our background study we found that there are two different methods of processing strategies (discussed in section 2.3.1). None of these approaches is better than the other, it is different style of processing strategies. Processing strategy depends on gender, the amount of information provided, cues, risk and complexity associated with the given task/problem.

In the post-task session, we used a standard questionnaire after slight modification (Appendix B) to answer this RQ. We summed up the responses of (7-point likert scale) first 5 questions for comprehensive score and the rest of the 5 responses for selective score. Maximum score of 35 for each participant. Outcome (Sum) from this analysis is shown in Figure 4.3 and summarized in table 4.3. We will describe results in the following subsection.

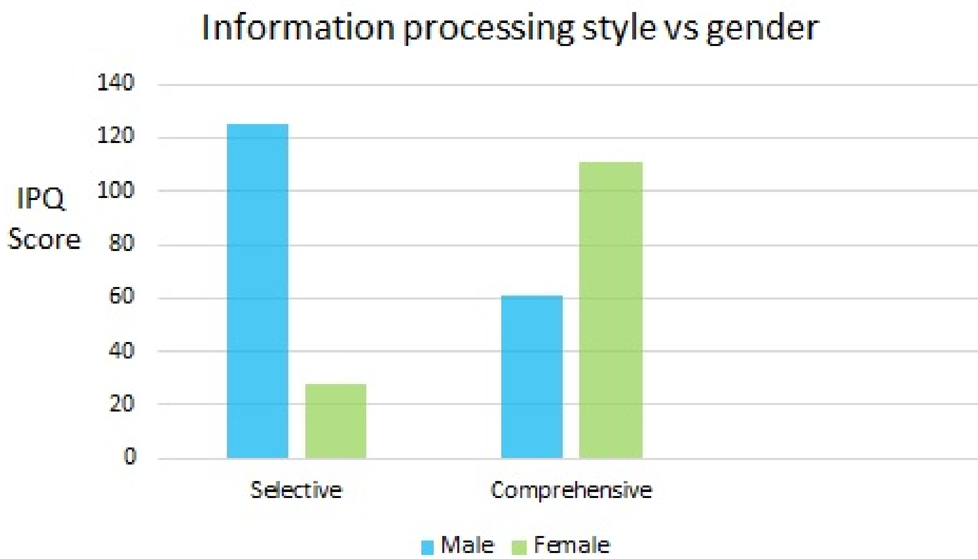

Figure 4.3: Information processing style Questionnaire (IPQ) score summary for using PR2 robot by different gender.

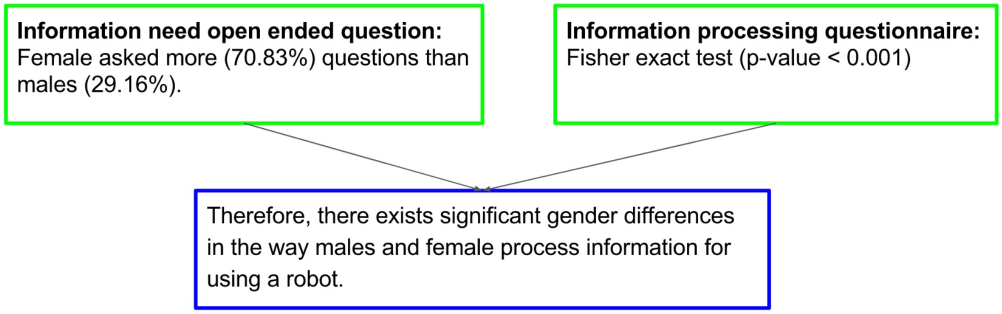

Figure 4.4: Information processing strategy for using PR2 robot result triangulation illustration.

4.2.2 Result triangulation: information processing style for using PR2 robot and gender

We found gender differences in information processing style in two ways:

1. First, by using Fisher exact test (two-tailed) on the data (sum) of table 4.3 we obtain p-value < 0.001.

2. Second, we can see from table 4.2 that female asked for more information (70.83%) than males (29.16%).

It is methodological triangulation [47, 53, 59, 74] because we obtained the same results using different methods (shown in Figure 4.4). They are:

1. Statistical analysis of post-task session questionnaire (Quantitative) and

2. Analysis of open ended question in the pre-task session (Qualitative), given in table 4.2.

Therefore, we reject the null hypothesis (H_0). And we conclude that females used comprehensive processing and males used systematic processing strategy for using the robot. This finding is congruent with selectivity hypothesis which says that males use a single cue that is most available to make a single inference. On the other hand, females takes consideration of all the details and do a piecemeal analysis to make an inference. This is also true in case of end user debugging strategies [10, 20]. Gender difference in selective information processing shown in table Figure 4.5.

Post-task session:

In the post-task session, we asked the following open ended question about information processing. *"How did you use that information?"*. In response users did not specify any strategy/process. This may be because we asked the question wrong. So we decided not to include these responses in our analysis.

In summary, we found significant gender difference in the way both gender process information using methodological triangulation.

4.3 Self-efficacy and gender in robot interaction (RQ3)

RQ3: ***Are there any gender differences in the self-efficacy that impact efficient use of the robot?***

Null Hypothesis (H_0): *There is no gender differences in the self-efficacy for using*

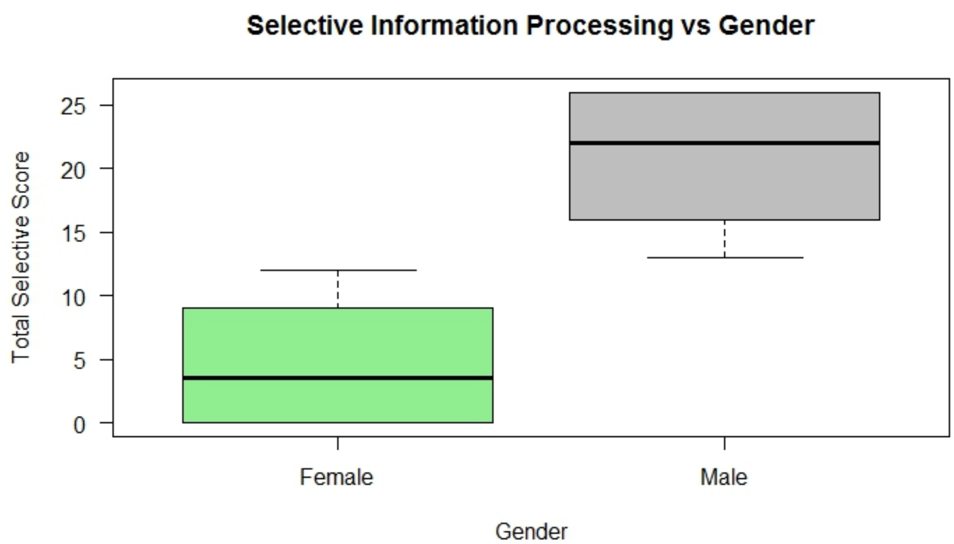

Figure 4.5: Gender difference in selective information processing while learning how to use PR2 robot.

a robot.

In section 3.6 we discussed how we obtained self-efficacy data by using 10-point likert scale standard questionnaire [26]. Complete questionnaire can be found in Appendix A. For each participant, responses from 10 questions were summed to self-efficacy score. Maximum score was 100 for each participant. After computing self-efficacy score for all the participants we performed statistical analysis on the rank transformed data with gender as the independent variable. Previous research showed gender difference in the self-efficacy in various computational problem solving instances such as math, debugging etc [7, 67] We applied ANOVA on our rank transformed self-efficacy data. One-way ANOVA (F (1,10) = 31.304, p-value = 0.00022927 ***) revealed significant gender difference in the self-efficacy for using PR2 robot. Thus we reject the null hypothesis and confirm that our female participants had lower self-efficacy than male participants for using the robot, see Figure 4.6.

ANCOVA analysis:

To understand whether there's any effect of previous video gaming experience (predictor variable) on self-efficacy (response variable) or not used Analysis of Covariance (ANCOVA). In this test gender served as the categorical variable. The result showed significant (p-value = 0.0226) effect on self-efficacy. Thus, we conclude that previous video gaming experience had an effect on the self-efficacy of male and females for using robot. This finding is similar to previous research of performance in debugging and introductory computer courses, which states that

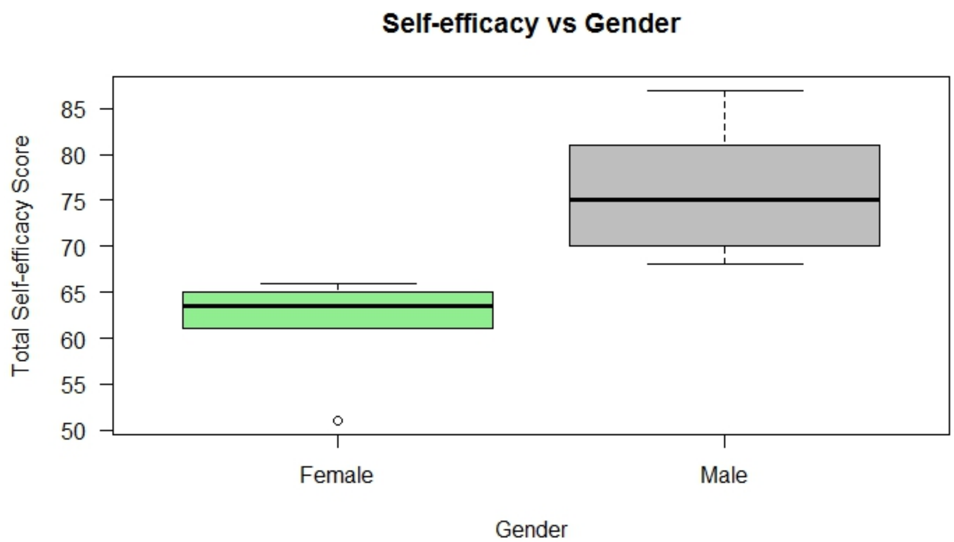

Figure 4.6: Gender difference in self-efficacy for using PR2 robot.

increasing experience may increase self-efficacy for females [7, 72]. This indicates that playing video games may have increased self-efficacy in the males for using a robot.

Non-parametric Spearman's rho test show this significant difference (rho = 0.664, p-value = 0.0184) with 95% confidence level.

Now we are going to discuss how self-efficacy relates to task-success for both male and female.

4.3.1 Self-efficacy, task success and gender

In the previous section, we discussed gender difference in self-efficacy for using a robot. Now we are going to discuss whether self-efficacy have an impact on task-performance or not. Previous research show that there exists a positive impact of self-efficacy on problem solving and improved performance [4, 6, 41].

Video data coding:

We code video data to understand how males and females performed in the given task (lets call it task score). Details about the task description can be found in section 3.7. We qualitatively code video data using three steps:

1. 100 is assigned for successfully completing the task.

2. Whenever after successful task execution participants made a mistake (dropping it after picking up correctly) we deduct 20.

3. We gave a 0 in case of participant failed the task completely.

4. 20 for partial task success.

We followed steps i) ii, iii) and iv) to code video data. Complete task coding can be found in Appendix D (Table D.7). Example of task success instances from video is shown in Figure 4.7 and Figure 4.8.

We performed ANOVA on the rank transformed data where task score is dependent variable and gender is independent variable. We found statistically significant gender difference in the task score (ANOVA $F(1,10) = 11.538$, $p =$

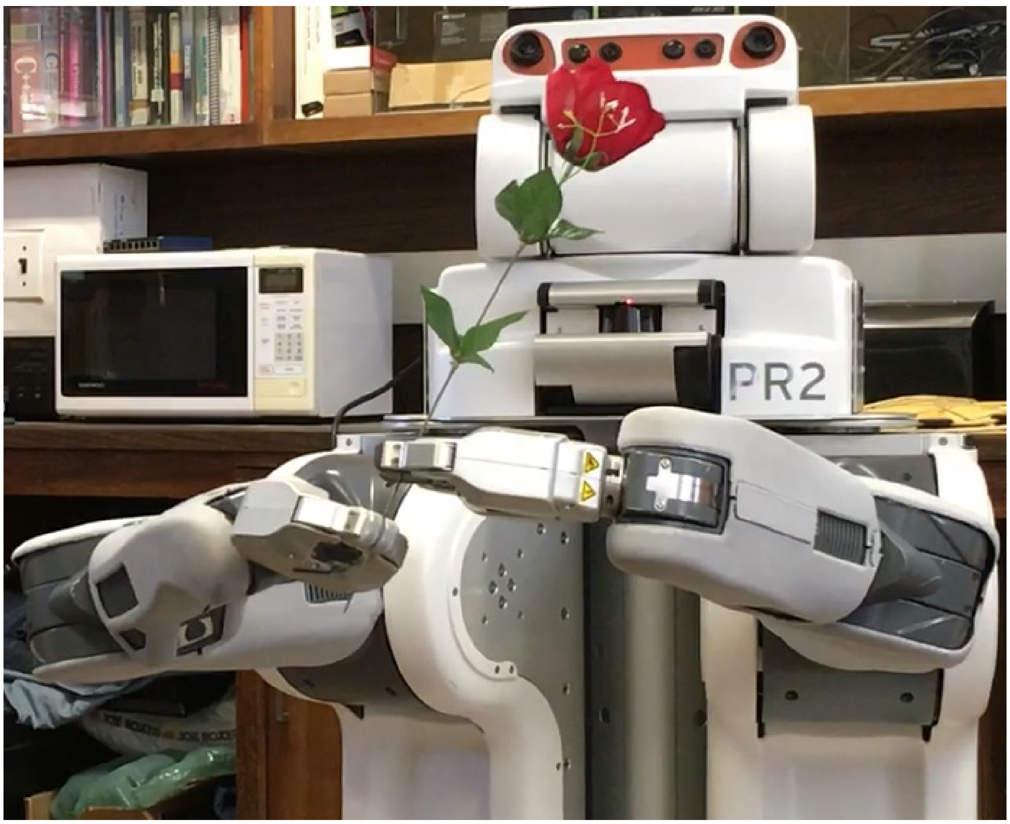

Figure 4.7: Task success example. Picking up a rose by its stem only and then passing it to to the other hand of PR2.

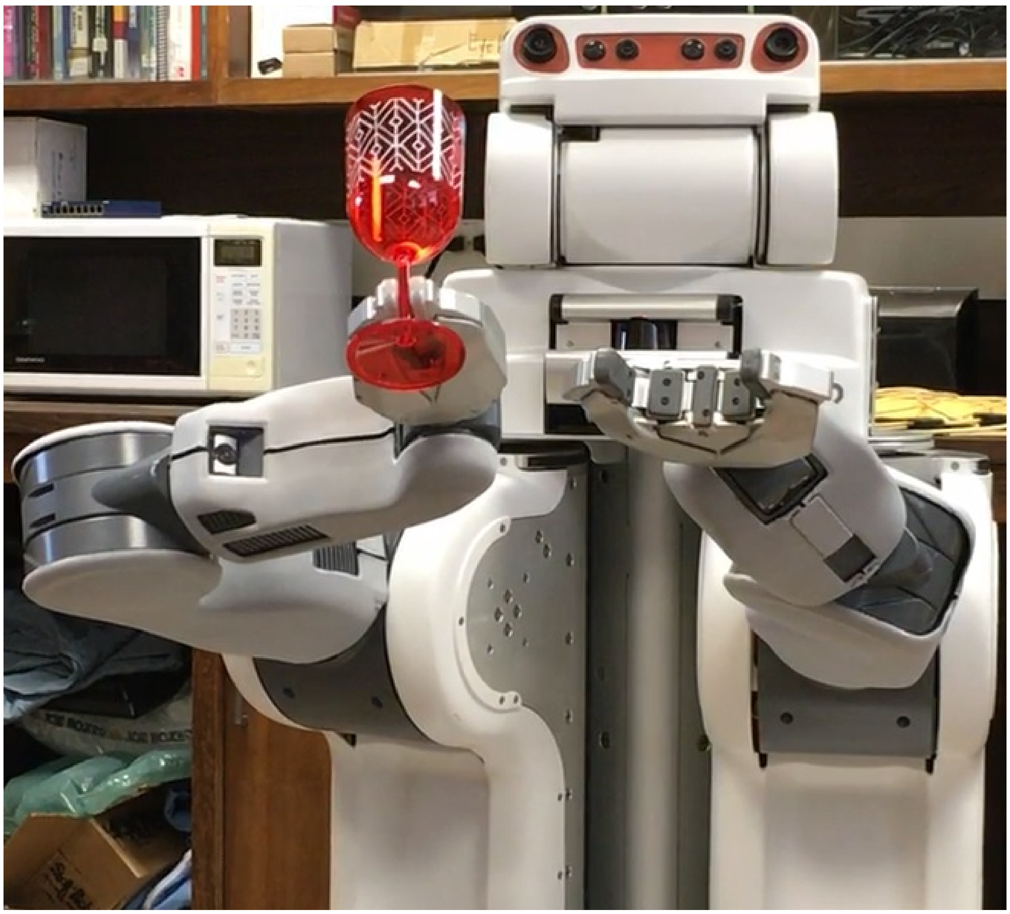

Figure 4.8: Task success example. Picking up a wine glass and have the PR2 drinking from the glass.

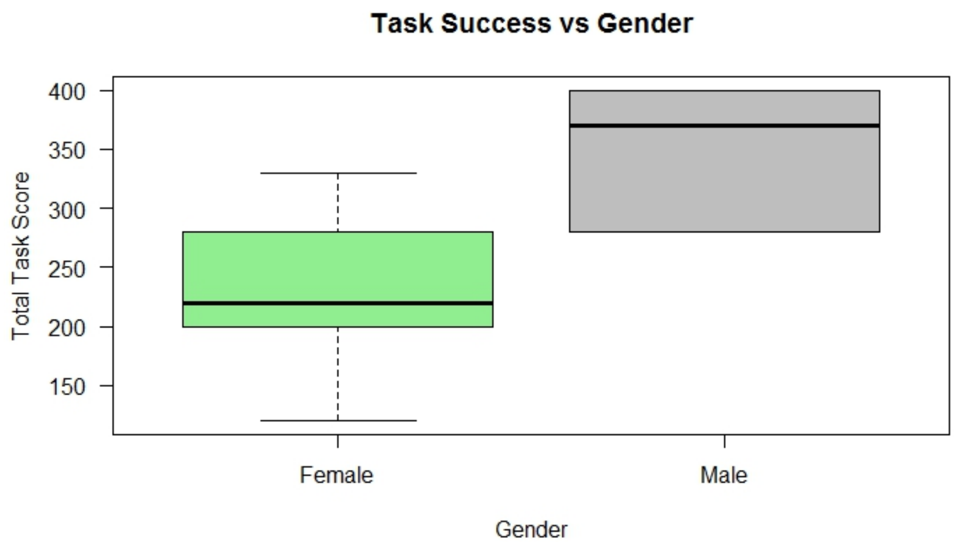

Figure 4.9: Gender difference in task success with PR2 robot.

0.0068072 **). This difference is shown in Figure 4.9. Males performed significantly better than females using the robot to perform various tasks.

ANCOVA analysis:

Previously we showed gender difference in self-efficacy and task score. Now we wanted to know how self-efficacy impacts task success across gender. Here we will test whether individuals having high self-efficacy result in higher task success or not as evident in previous researches.

1. We performed ANCOVA test with task score as dependent variable, video gaming experience as covariate and gender as a factor. This test showed no significant interaction (p-value = 0.09).

2. We took task score as dependent variable, self-efficacy as independent variable and gender as factor. ANOVA revealed no significant difference (p-value = 0.13).

3. We test whether task success is influenced by self-efficacy, video gaming experience and gender. MANOVA revealed no significant difference (p-value = 0.74).

Furthermore, we test for correlation coefficient between task score and self-efficacy. We performed spearman's rho test (p-value = 0.132). There was no statistical basis indicating that task success was significantly impacted by self-efficacy and video gaming experience.

4.3.2 Self efficacy, task time and gender

Measuring task time:

Total task completion time (in minutes) was computed by summing up the time taken for completing each task (task description can be found in table 3.2). This was done for all participants. By completing task we mean when users decided that they have completed (consists of successful/failure attempts) the task. Data showed that on average males (12.45 minutes) took less time than females (21.56 minutes) to complete the tasks. Task completion time coding can be found in Appendix D (Table D.8).

ANCOVA analysis:

Gender	*Self-efficacy score*	*Task score*	*Total time (Avg min)*	*Tinkering count*
Females	61.67 (avg)	228.3	21.56	29
Males	76 (avg)	350	12.45	46

Table 4.4: Self-efficacy score, total task score and task time versus gender while using the PR2 robot.

1. Video gaming experience showed significant interaction (F-value (1,10) = 6.61, p-value = 0.030) on task time with gender as factor.

2. ANOVA analysis on task time (dependent), self-efficacy, and gender (factor) revealed significant difference (F-value (1,10) = 5.16, p-value = 0.049).

For further verification, we test for correlation coefficient between task time and self-efficacy. But there was no significant (p-value = 0.2697) correlation between them. Also there was no significant correlation between task time and video gaming experience (p-value = 0.095). These results show that task time is influenced by self-efficacy and video gaming experience. This may be caused by the fact that users who play video games are probably comfortable using joysticks (as mentioned by users) which may reduce task time. Self-efficacy, task score, task time is given in Table 4.4.

4.4 Tinkering and gender in robot interaction (RQ4)

RQ4: ***Are there any gender differences about tinkering with the robot?***

Null Hypothesis (H_0): *There is no gender differences about the tinkering with a robot.*

We will discuss how we code for tinkering from pre-task and task session and analyzed the data in the following sections.

4.4.1 Measuring tinkering from pre-task session

To understand tinkering from gender perspective, we asked participants the following open ended question before doing anything with the robot to understand their mental models.

"*How often do you play with a new hardware device? Why? Why not?*"

Responses were qualitatively coded by the researcher. A few responses are given in Table 4.5. We code for i) playful: when a response contains keywords as "fun","play","enjoy", "tinker", "love trying" etc. ii) not-playful: when a response demonstrates a negative approach about playfulness towards a new device (exp. "not much", "not often"). Data and associated codes can be found in Appendix D (Table D.1). After categorizing all the responses, we applied statistical tests to analyze our data and validate our hypothesis. We discover that, males exhibit more of a playful mindset when it comes to using a new hardware device. On the other hand, females exhibit more of a negative attitude towards playing with new hardware devices. Fisher exact test on our data indicates (p-value = 0.0021**) this difference. Statistical results indicates that there exists significant gender difference when it comes to playing with new hardware devices, thus we reject the null hypothesis (H_0). This finding is congruent with previous researches of tinkering on the effectiveness of debugging software, in scientific education, learning and

Gender	*Quotes*
Male	"I love trying new devices such as VR, tablets, smart watches."
Male	"Often, because it is fun."
Female	"I dont play with new hardware device, I am a peoples person."
Female	"Almost never. I am not tech savvy at all."
Female	"I dont have a lot of money."
Male	"I love to play with all new gadgets come up in the market."

Table 4.5: Response from open ended question about tinkering across genders.

development [9, 52, 69].

4.4.2 Measuring tinkering from task session with PR2 robot

In our research, we defined tinkering as "playful experimentation" with the robot. We allowed participants to use three basic controls (open/close fingers, moving hands, arm movements), which in turn provides users limited features to tinker with. Therefore, its not what features users chose to use, instead the different ways users completed given tasks is termed as tinkering.

In this part of data analysis, we code video data for the task session (task session described in section 3.7). To compute "***tinkering count***" we used the following definition:

"whenever participant use a ***different approach*** to complete the task using robot" is defined as an instance of tinkering. We call an approach different if users use any of the following:

1. Using one robot hand (left or right) vs both hands.

2. Picking object from top/bottom/side (in case of tall objects for example).
3. Robot claw position horizontally/vertically for left/right/both hand (for objects with small/no height).
4. picking up objects using its tip/body/handle/stem.

For example, Figure 4.10 and 4.11 shows a participant trying picking up a bowl from side and top respectively. The participant (shown in Figure 4.10) realized that he/she chose a wrong approach and then corrected his/her approach for the task (shown in Figure 4.11). We count for distinct tinkering instances. Since the robot has only 2 fingers compared to robots having 5 fingers, the number of choices users had was fairly small. More tinkering examples can be found in Figure 4.12. Figure 4.13 shows an example where a participant was not tinkering, in other words changing his/her approach to pick up the wine glass.

We simply count those instances from video data, where the audio gives context of that tinkering event. We then summed up these events across all the tasks for each participant, this gives us tinkering score. Codes can be found in Appendix D (Table D.2). Coded data can be found in Appendix D (Table D.3). The coded data is then statistically analyzed and plotted (shown in Figure 4.14). We found that there exists gender difference in how much male and female tinker with the robot, so we reject null hypothesis (H_0). We applied ANOVA on our rank transformed data. One-way ANOVA ($F(1,10) = 13.158$, p-value $= 0.0046326$ **) revealed significant gender difference while tinkering with robot.

Figure 4.10: Tinkering example. Picking up a bowl from the side of the bowl.

Figure 4.11: Tinkering example. Picking up a bowl by the tip of the bowl.

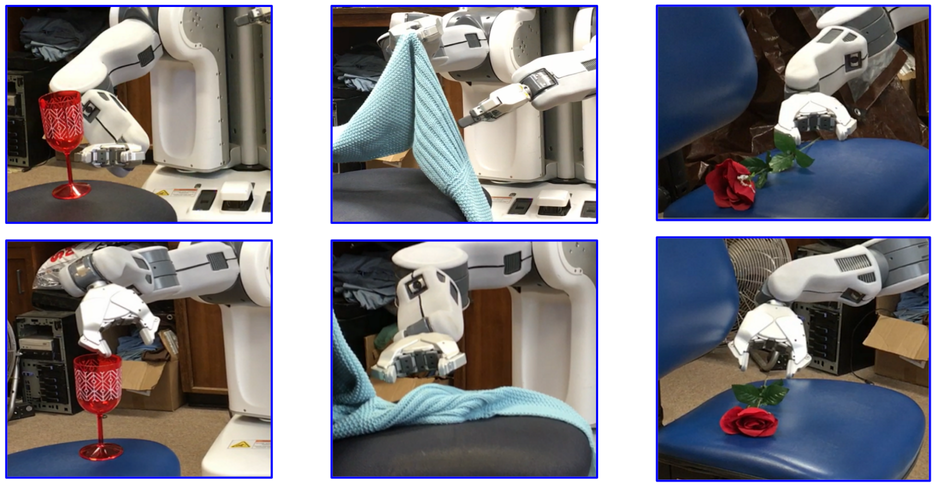

Figure 4.12: Tinkering while grasping wine glass (Object-stem-left, Object-tip-left), scarf (claw-both, claw-horizontal-left), rose (claw,horizontal-left, claw-vertical-left).

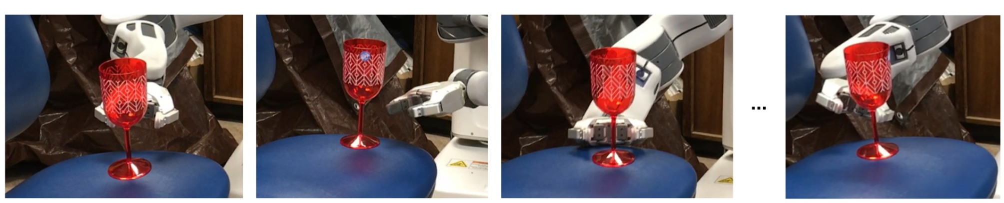

Figure 4.13: An example where a participant was not tinkering at all.

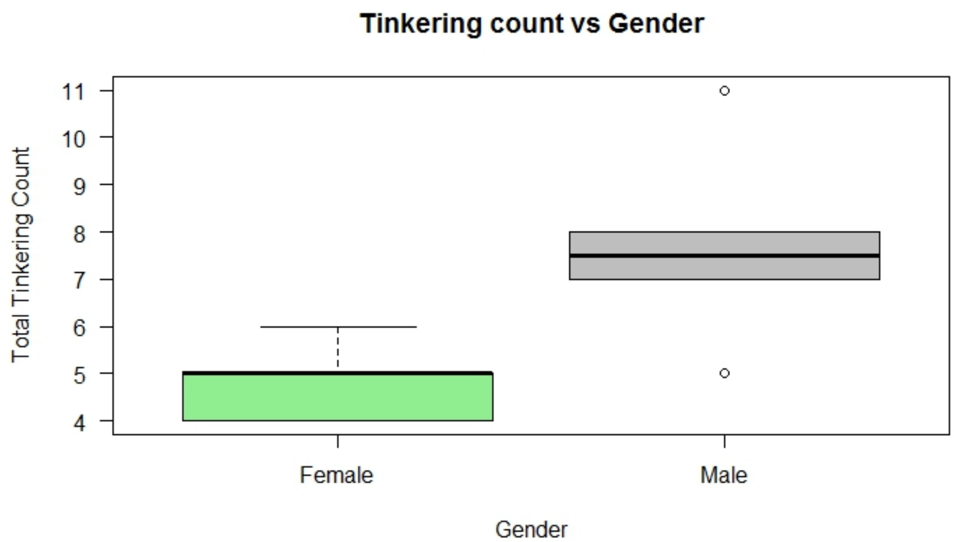

Figure 4.14: Gender difference in tinkering with the PR2 robot for grasping task.

4.4.3 Result triangulation: tinkering in HRI and gender

It is a good practice to cross validate or use more than one source of data to confirm findings in any research study. In this subsection, we are going to summarize our research findings of tinkering while using robot. In section 4.4.1 we discussed our finding from open ended question in the pre-task session, and in section 4.4.2 we discussed tinkering in the task session. Two different sources result in congruent result. This is a form of methodological triangulation because (shown in Figure 4.15) we obtain same result using different methods [47, 59, 53, 74]. The methods are:

1. Analysis of open ended question (qualitative) and

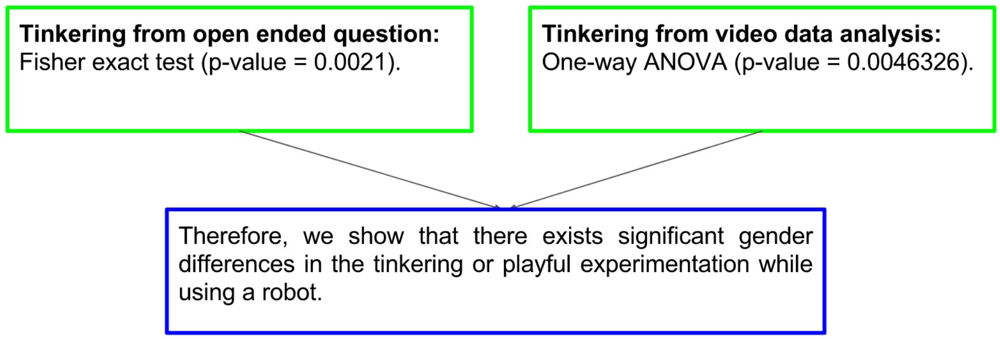

Figure 4.15: Tinkering with PR2 result triangulation illustration.

2. Analysis of tinkering count (quantitative) from observation of video data.

Therefore, we reject null hypothesis (H_0) and conclude that there exist gender difference in tinkering while using a robot that users have never used before.

Non parametric spearman's correlation coefficient test showed a significant positive correlation of 0.7345532 between tinkering and video gaming experience having (S = 75.918 and p-value = 0.0065) with 95% confidence interval.

ANCOVA analysis:

We performed Analysis of Covariants (ANCOVA) with tinkering as the dependent variable, gender as independent variable and video gaming experience as covariate. In this case video gaming experience had significant (p-value <2e-16 ***) effect on tinkering.

4.4.4 Task success, tinkering and gender in robot interaction

We are going to discuss several test results that we performed to determine relationship among task success, tinkering and gender.

1. Non parametric spearman's rho test also reveal a positive significant correlation 0.7319251 between task score and tinkering (S = 76.669 and p-value = 0.006807).

2. Then by fitting a linear model between task score as dependent variable and tinkering (from open ended question) as independent variable we found F(1,10) = 10.71, $R^2 = 0.517$, p = 0.008404.

3. We fit a linear model on tinkering count (from task session), this test showed a significant correlation (F(1,10) = 4.645, $R^2 = 0.317$, p-value: 0.05). This result is shown in Figure 4.16.

ANCOVA analysis:

1. We performed ANOVA with task score as dependent variable, tinkering as independent variable and gender as a factor. The result indicates p-value = 0.0084. Therefore tinkering may have impact on task success. Males tinker more, so this could be a possible explanation for their high task success comparing to females.

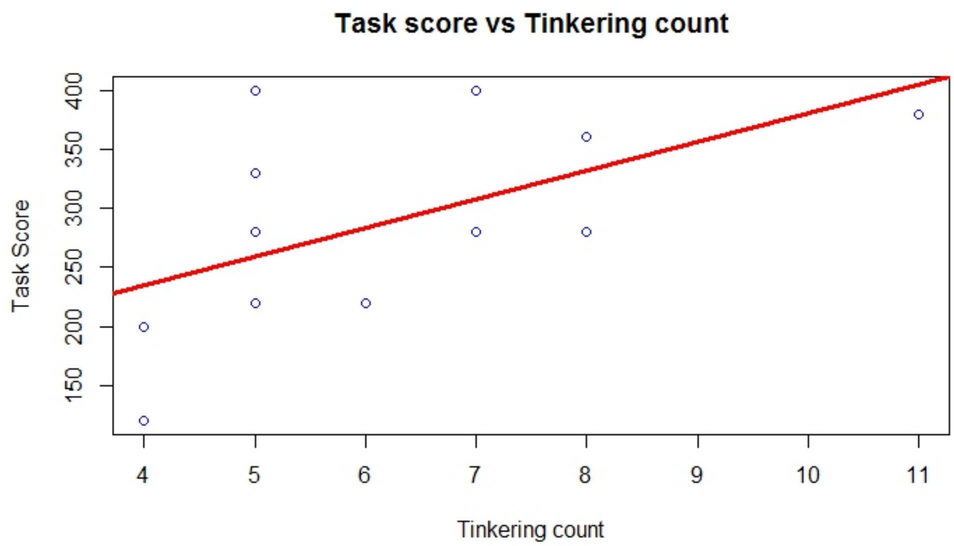

Figure 4.16: Task success using PR2 and how it relates to tinkering count.

2. We then checked for a combination of video gaming experience, tinkering, gender as independent variable and their impact on task success (dependent variable). ANCOVA indicated no significant difference in this case.

A possible explanation for positive correlation between task success and tinkering could be that users who tinkered more possibly changed their approach to the task quite often. In the video, we observed that males changed their approach very frequently to complete the task, i.e. they decided to move on to a new strategy when a previous strategy failed. This lead to males greater task success than females (who tinkered less).

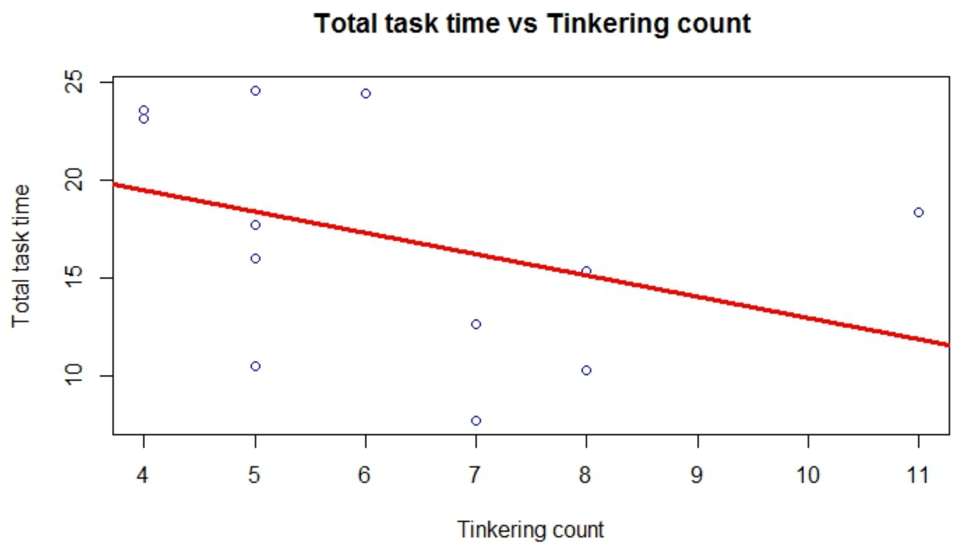

Figure 4.17: Total task completion time how it relates to tinkering count illustration.

4.4.5 Task time, tinkering and gender in robot interaction

1. Non parametric Spearman's rho test also reveal negative -0.7724873 correlation between task time and tinkering (S = 506.93 and p-value = 0.0032) with 95% confidence interval.

2. Then we fit a linear model on task time as dependent and tinkering as independent variable, we obtained F(1,10) = 17.21, R^2 = 0.6324, p = 0.001986. This correlation is shown in Figure 4.17.

ANCOVA analysis:

We computed ANCOVA on task time as dependent variable, video gaming experience as covariate, tinkering as independent variable and gender as factor. This test revealed significant interaction between tinkering and task time which we mentioned earlier.

Total tinkering count, task score, task time is given in Table 4.4. Negative correlation between task completion time and tinkering can be explained by the fact that males changed their approach to a task more frequently than females (observation from video). They took less time in deciding if an approach did not work, they switched fast to a different approach, that's why males took less time than females to complete given task.

Chapter 5: Discussion and Threats to Validity

From our background research we have come to know that there has been significant amount of research that has been done to understand users information need, information processing style, self-efficacy, tinkering in education, social psychology and computer related problem solving. These studies report gender differences in these factors. But there is almost no research that tried to understand these factors in human robot interaction (HRI). In this research, we took this initiative to broaden our understanding of the aforementioned factors in HRI domain by examining users operating a humanoid robot from distance (robot teleoperation) from gender perspective. Previous researches have looked into gendered humanoid robots in persuasion, eliciting information, social facilitation and how they influence different gender (males and females). All these findings are significant from because of the future application of robots (discussed in section 2.2).

We discovered several significant findings from this research study. We tested our results with both parametric and non-parametric tests. Parametric test showed similar results to non-parametric test (statistical significance) but we report only non-parametric test result in our final analysis. We performed our non-parametric tests by first rank transforming our data and then applying ANOVA. We will now briefly discuss our research findings.

Gender difference in information need and self-efficacy are validated using non-parametric ANOVA test. Gender difference in information processing style and tinkering has been cross validated by methodological triangulation (both qualitative and quantitative). Correlation between video gaming experience and self-efficacy, video gaming experience and tinkering, tinkering and task score, tinkering and task time are tested using correlation tests (non-parametric Spearman's rho test). Moreover, we had equal number of males and females in our study that eliminates errors due to unequal sample size.

In this study, we found that female participants asked for more information (strategy, feature, functionality) than male participants (strategy) for using the robot. This goes back to the research of Meyers levy's [62] processing strategies where she mentioned that females engage in more detailed elaborative processing of all available cues, and males use of heuristic based processing. It is statistically significant that females asked for more information than males on how to use the robot both by the number of categories of questions asked and the number of questions asked.

In our study, females processed information for using the PR2 robot comprehensively (gather, gather, gather, ... do, do, do) and males used systematic processing (gather .. do, gather .. do, gather .. do). Background research in processing strategies indicates that females are more concerned about others, but males are more goal (self) oriented. This can be backed up using evidence from the environment where we were brought up. Females are generally brought up in a structured environment (where they are given instruction) and males are often grow in an unstructured environment (where they learn to grapple). Parenting, mother's and father's attitude towards their girls and boys often plays a role in how different gender processes information. All these evidence contributes to females include and covering everything style of processing, and males select what looks best at the moment of selective processing [83].

Our results show that males had greater self efficacy than females when it comes to complete the given tasks using robot. Males had more video gaming experience. And video gaming experience is correlated with self efficacy (positively). ANCOVA test indicates self-efficacy and video gaming experience might influence task time. We found no relation between self-efficacy and task score.

We found that males in our study tinkered more often than female. This finding is congruent with previous research about tinkering in software literature and education. From video we observed that males tend to frequently changed their approach to a given task if a previous approach did not work in the first place. This probably lead them to higher success and complete task in less time. Video gaming experience is correlated with tinkering. We found tinkering is correlated with task score (positively) and task completion time (negatively).

We wanted to measure participants attitudes towards risk by asking a 10-point liker scale type question in the post-task session (Appendix B). The question was:

Q-2 "How often did you make any risky move that you were not sure of?" 1 2 3 4 5 6 7 8 9 10

We computed summation of these responses (male = 35, females = 32). Rank transformed ANOVA showed no significant ($F(1,10) = 0.096502$, p-value = 0.76245) gender differences for attitudes towards risk with a robot. This can be explained by the fact that we actually gave no task that actually requires risk taking, thus measuring it was almost impossible.

For measuring motivation for using the robot we asked an open ended question in the pre-task session (Appendix A). The question was:
"Suppose you are given a good sum of money, would you buy this robot? Mention why/why not?"
After qualitative coding of the responses, we found that 7 participants showed motivation for buying the robot for various reasons (clean bed, doing repetitive tasks etc.). 3 participants response was a clear no. And 2 participants responded yes/no, meaning that their response was conditional. So we could not come up with a conclusive result for motivation.

As with any qualitative research, our study is not free from threats. The following section describes potential threats that can invalidate our research findings.

5.1 Threats to Validity

Validity of any research indicates whether we are measuring the right things to answer our research questions. No empirical study is free from threats. In the following subsections, we are going to discuss internal, external, construct and conclusion threats to our study.

5.1.1 External Validity

External validity refers to the fact that whether findings from our experiment is generalizable or not [15, 31, 74, 89]. To avoid demographic bias we made our study open to all. But eventually only students reached out and showed interest to participate in our study. Thus, our result may not be generalizable across different demographics due to small sample size of only 12 participants (6 males and 6 females). The results may not be representative of different age groups as well. Moreover, there could be bias associated with researcher gender. We select a mixed of easy and complex tasks, complexity of the task may lead to unprecedented bias.

5.1.2 Internal Validity

It refers to the risks that may hinder the causal relationship between dependent and independent variables [15, 31, 74, 89]. Our participants had no previous experience with robots, but they had experience with video games. Males and females had significantly different video gaming experience. To eliminate the chance that video gaming experience might influence our experiment result, we used video gaming experience as covariates in all ANCOVA analysis. We also test our result with multiple statistical methods and methodological triangulation but we only report on non-parametric results. Only after cross validation we reported our research finding.

5.1.3 Construct Validity

It is associated with the risk that whether the theory and the operational measures taken reflect the mind set of the researcher and what is being asked in the research questions [15, 31, 74, 89]. To measure self-efficacy and information processing style we used pre-existing standard questionnaire from literature after slight modification (with no modification of scaling). Information need is measured from response to an open ended question. For measuring tinkering we asked open ended question and used observation of video data. This eliminates the probability of having construct threat to our study.

5.1.4 Statistical Conclusion Validity

Conclusion validity refers to the threat associated with the incorrect assumptions before we statistically test a relationship between two variables [15, 31, 74, 89]. Our sample size was small but had equal number of males and females. This avoids problems associated with unequal sample size. We rank transformed our data and then applied ANOVA. We test each relationship with multiple statistical tests and triangulated results. We reported correlation between tinker, task success and task time after multiple correlation test reports its significance. Moreover, we used co-variates in all our ANCOVA's. Every result is tested using multiple approaches and methodological triangulation which certainly validates our findings.

Chapter 6: Conclusions and Future Work

In this chapter, we are going to summarize findings from our empirical research based investigation. We used both qualitative and quantitative data analysis methodologies. We will discuss our research findings and future scope in two separate sections.

6.1 Conclusion

In this study we tried to understand users information need, information processing style, self-efficacy and tinkering, and how they influence task success and task completion time using the PR2 robot that they have never used before from gender perspective. Results from our research is summarized as below.

1. ***RQ1:*** We found that information required for task completion for using robot falls under three categories. They are feature, strategy, and functionality. This finding is useful from the robot design perspective. Our result indicate significant (statistically) gender difference about the amount of information needed on how to use a robot (both by number of categories of questions and number of questions). Our study results show that females asked for more information compared to males. Males wanted to know about "strategy" and "feature" related information, whereas females asked for "fea-

ture","functionality" as well as "strategy" related information before using the robot. This finding is congruent with social cognition literature that dictates that as the basis of judgment males rely on cues that are highly available and females relied more on all the available cues [82].

2. ***RQ2:*** We found statistically significant gender differences in the way males (selectively) and females (comprehensively) processed information while performing object manipulation task using PR2 robot. This finding is consistent with existing literature of processing strategies. We validated this outcome using methodological triangulation. First, using quantitative analysis of post task questionnaire (10-point likert scale) responses and then qualitatively analyzing open ended question responses in the pre-task session.

3. ***RQ3:*** Difference in the self-efficacy score in our study result indicates that male participants were more confident about manipulating the robot than our female participants. They took less time, achieved greater task success. This finding is consistent with self-efficacy literature in math education and computer problem solving such as debugging. Bandura et. al. reports that people with greater self-efficacy result in task persistence, amount of effort put and result in higher task performance [4]. We found no significant correlation between self-efficacy with task success, and self-efficacy and task time. ANOVA indicates self-efficacy might influence task time. Spearman's non-parametric test indicates that video gaming experience and self-efficacy are strongly correlated.

4. ***RQ4:*** We found that there exist significant gender difference in the willingness to play or tinker with a new hardware device. Males are more willing to play with new hardware device than females. ANOVA test result indicate that playfulness probably have result in males to perform better in given task than females. This has been confirmed by Pearson and spearman's correlation test. There exist a positive correlation between video gaming experience and tinkering, but video gaming experience has no correlation with task success or task time. Tinkering is positively correlated with task success and negatively correlated with task time. Therefore, increase in tinkering may have result in higher task success and reducing task completion time with the robot.

6.2 Future work and implications

We have the following research implication and future opportunities.

- Grasping is still considered a hard problem in human robot interaction. That is one reason that the outcome from this research can be useful for building intelligent user interfaces that provides the right information need at the right time.

- Signifier and affordance is an important usability concept. Unfortunately, not all design support them. This makes a design usable for one group while not usable by the other. We already discussed that using GenderMag personas helped companies to fix inaccessibility issue within their software.

Similarly, persona can be tested for usability of robots. This is an open research direction.

- Knowledge gained from this study can benefit design of robots that are gender inclusive, i.e. accessible to a wider range of population that enable greater user experience.

- This research clarifies the fact that males and females use robot differently. Before robots make their way to our homes, we need to make sure that whether we are making then with sufficient user concern from useful and usable point of view. Testing robots with different groups of people (adults, senior citizens etc.) might help improve its acceptance at a faster pace than it is right now.

- In future, we would like to extend our study by incorporating more complex tasks with wide range of different population. Furthermore, robots gender and how it influence males and females task behavior would be interesting to research as well.

- We did not find any conclusive outcome in risk taking probably because we asked/phrased the question wrong or with not enough information which provides better context. This is an open research scope. The same goes for motivation as well.

Bibliography

[1] John Baer and James C Kaufman. Gender differences in creativity. *The Journal of Creative Behavior*, 42(2):75–105, 2008.

[2] Ravi Balasubramanian, Ling Xu, Peter D Brook, Joshua R Smith, and Yoky Matsuoka. Human-guided grasp measures improve grasp robustness on physical robot. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 2294–2301. IEEE, 2010.

[3] Ravi Balasubramanian, Ling Xu, Peter D Brook, Joshua R Smith, and Yoky Matsuoka. Physical human interactive guidance: Identifying grasping principles from human-planned grasps. *IEEE Transactions on Robotics*, 28(4):899–910, 2012.

[4] Albert Bandura. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191, 1977.

[5] Albert Bandura. Human agency in social cognitive theory. *American psychologist*, 44(9):1175, 1989.

[6] Julian Barling and Russell Beattie. Self-efficacy beliefs and sales performance. *Journal of Organizational Behavior Management*, 5(1):41–51, 1983.

[7] Laura Beckwith, Margaret Burnett, Susan Wiedenbeck, Curtis Cook, Shraddha Sorte, and Michelle Hastings. Effectiveness of end-user debugging software features: Are there gender issues? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 869–878. ACM, 2005.

[8] Laura Beckwith, Derek Inman, Kyle Rector, and Margaret Burnett. On to the real world: Gender and self-efficacy in excel. In *IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2007)*, pages 119–126. IEEE, 2007.

[9] Laura Beckwith, Cory Kissinger, Margaret Burnett, Susan Wiedenbeck, Joseph Lawrance, Alan Blackwell, and Curtis Cook. Tinkering and gender in end-user programmers' debugging. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 231–240. ACM, 2006.

[10] Laura A Beckwith, Valentina I Grigoreanu, Neeraja Subrahmaniyan, Susan Wiedenbeck, Margaret Burnett, Curtis R Cook, Karin Bucht, Russell Drummond, et al. Gender differences in end-user debugging strategies. Technical report, Corvallis, OR: Oregon State University, Dept. of Computer Science, 2007.

[11] Charles Bell. *Beginning sensor networks with Arduino and Raspberry Pi.* Apress, 2014.

[12] Susan Bergin and Ronan Reilly. The influence of motivation and comfort-level on learning to program. In *Proceedings of the PPIG*, volume 17, pages 293–304. Citeseer, 2005.

[13] Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan. Autonomous vehicles need experimental ethics: are we ready for utilitarian cars? *arXiv preprint arXiv:1510.03346*, 2015.

[14] Cynthia Breazeal. Toward sociable robots. *Robotics and autonomous systems*, 42(3):167–175, 2003.

[15] Marilynn B Brewer, H Reis, and C Judd. Research design and issues of validity. *Handbook of research methods in social and personality psychology*, pages 3–16, 2000.

[16] Elizabeth Broadbent, I Han Kuo, Yong In Lee, Joel Rabindran, Ngaire Kerse, Rebecca Stafford, and Bruce A MacDonald. Attitudes and reactions to a healthcare robot. *Telemedicine and e-Health*, 16(5):608–613, 2010.

[17] Margaret Burnett, Scott D Fleming, Shamsi Iqbal, Gina Venolia, Vidya Rajaram, Umer Farooq, Valentina Grigoreanu, and Mary Czerwinski. Gender differences and programming environments: across programming populations. In *Proceedings of the 2010 ACM-IEEE international symposium on empirical software engineering and measurement*, page 28. ACM, 2010.

[18] Margaret Burnett, Anicia Peters, Charles Hill, and Noha Elarief. Finding gender-inclusiveness software issues with gendermag: a field investigation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2586–2598. ACM, 2016.

[19] James P Byrnes, David C Miller, and William D Schafer. Gender differences in risk taking: A meta-analysis., 1999.

[20] Patricia Cafferata and Alice M Tybout. Gender differences in information processing: A selectivity interpretation, cognitive and affective responses to advertising, 1989.

[21] Daniel J Canary and Kimberley S Hause. Is there any reason to research sex differences in communication? *Communication Quarterly*, 41(2):129–144, 1993.

[22] Adriano Cavalcanti. Robots in surgery. In *Plenary Lecture, Euro Nano Forum*, 2005.

[23] C CDATA-Nass, Youngme Moon, and Nancy Green. Are computers gender-neutral? gender stereotypic responses to computers. *Journal of Applied Social Psychology*, 27(10):864–876, 1997.

[24] Gary Charness and Uri Gneezy. Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1):50–58, 2012.

[25] Tiffany L Chen, Matei Ciocarlie, Steve Cousins, Phillip M Grice, Kelsey Hawkins, Kaijen Hsiao, Charles C Kemp, Chih-Hung King, Daniel A Lazewatsky, Adam E Leeper, et al. Robots for humanity: Using assistive robotics to empower people with disabilities. *IEEE Robotics & Automation Magazine*, 20(1):30–39, 2013.

[26] Deborah R Compeau and Christopher A Higgins. Computer self-efficacy: Development of a measure and initial test. *MIS quarterly*, pages 189–211, 1995.

[27] Judd Cramer and Alan B Krueger. Disruptive change in the taxi business: The case of uber. *The American Economic Review*, 106(5):177–182, 2016.

[28] William K Darley and Robert E Smith. Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. *Journal of Advertising*, 24(1):41–56, 1995.

[29] Kerstin Dautenhahn, Michael Walters, Sarah Woods, Kheng Lee Koay, Chrystopher L Nehaniv, A Sisbot, Rachid Alami, and Thierry Siméon. How may i serve you?: a robot companion approaching a seated person in a helping context. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 172–179. ACM, 2006.

[30] Tamara Denning, Cynthia Matuszek, Karl Koscher, Joshua R Smith, and Tadayoshi Kohno. A spotlight on security and privacy risks with future household robots: attacks and lessons. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 105–114. ACM, 2009.

[31] Ellen A Drost. Validity and reliability in social science research. *Education Research and Perspectives*, 38(1):105, 2011.

[32] Clemens Eppner and Oliver Brock. Planning grasp strategies that exploit environmental constraints. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 4947–4952. IEEE, 2015.

[33] Anne Fausto-Sterling. Myths of gender: biological theories about man and woman, 1985.

[34] Matthew Feeney and Rideshare companies Uber. Is ridesharing safe? *Cato Policy Analysis*, 767:2, 2015.

[35] Alan Feingold. Gender differences in personality: a meta-analysis. *Psychological bulletin*, 116(3):429, 1994.

[36] Bernd Figner and Elke U Weber. Who takes risks when and why? determinants of risk taking. *Current Directions in Psychological Science*, 20(4):211–216, 2011.

[37] T Fong, C Thorpe, and C Baur. Collaboration, dialogue and human-robot interaction, 10th international sumposium of robotics research (lorne, victoria, australia). In *Proceedings of the 10th International Symposium of Robotics Research*, 2001.

[38] Willow garage. Pr2 overview. http://www.willowgarage.com/pages/pr2/overview, March 19, 2017. Accessed: 2017-04-27.

[39] Valentina Grigoreanu, Margaret Burnett, Susan Wiedenbeck, Jill Cao, Kyle Rector, and Irwin Kwan. End-user debugging strategies: A sensemaking perspective. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 19(1):5, 2012.

[40] Valentina Grigoreanu, Jill Cao, Todd Kulesza, Christopher Bogart, Kyle Rector, Margaret Burnett, and Susan Wiedenbeck. Can feature design reduce the gender gap in end-user software development environments? In *Visual*

Languages and Human-Centric Computing, 2008. VL/HCC 2008. IEEE Symposium On, pages 149–156. IEEE, 2008.

[41] Gail Hackett and Nancy E Betz. An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for research in Mathematics Education*, pages 261–273, 1989.

[42] Christine R Harris, Michael Jenkins, and Dale Glaser. Gender differences in risk assessment: why do women take fewer risks than men? *Judgment and Decision making*, 1(1):48, 2006.

[43] Charles Hill, Shannon Ernst, Alannah Oleson, Amber Horvath, and Margaret Burnett. Gendermag experiences in the field: The whole, the parts, and the workload. In *Visual Languages and Human-Centric Computing (VL/HCC), 2016 IEEE Symposium on*, pages 199–207. IEEE, 2016.

[44] Daniel Howard and Danielle Dai. Public perceptions of self-driving cars: The case of berkeley, california. In *Transportation Research Board 93rd Annual Meeting*, number 14-4502, 2014.

[45] Riitta Jääskeläinen. Think-aloud protocol. *Handbook of translation studies*, 1:371–373, 2010.

[46] Beverley Jane. Online reflections about tinkering in early childhood: A socio-cultural analysis. In *Asia-Pacific Forum on Science Learning and Teaching*, volume 7, page n1. ERIC, 2006.

[47] Todd D Jick. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly*, 24(4):602–611, 1979.

[48] Frank R Kardes, Maria L Cronley, James J Kellaris, and Steven S Posavac. The role of selective information processing in price-quality inference. *Journal of Consumer Research*, 31(2):368–374, 2004.

[49] Cory David Kidd. *Sociable robots: The role of presence and task in human-robot interaction*. PhD thesis, Massachusetts Institute of Technology, 2003.

[50] Cory Kissinger, Margaret Burnett, Simone Stumpf, Neeraja Subrahmaniyan, Laura Beckwith, Sherry Yang, and Mary Beth Rosson. Supporting end-user debugging: what do users want to know? In *Proceedings of the working conference on Advanced visual interfaces*, pages 135–142. ACM, 2006.

[51] I Han Kuo, Joel Marcus Rabindran, Elizabeth Broadbent, Yong In Lee, Ngaire Kerse, RMQ Stafford, and Bruce A MacDonald. Age and gender factors in user acceptance of healthcare robots. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 214–219. IEEE, 2009.

[52] Maarten H Lamers, Fons J Verbeek, and Peter WH van der Putten. Tinkering in scientific education. In *Advances in computer entertainment*, pages 568–571. Springer, 2013.

[53] Nancy L Leech and Anthony J Onwuegbuzie. An array of qualitative data analysis tools: A call for data analysis triangulation. *School psychology quarterly*, 22(4):557, 2007.

[54] Mark R Lepper. Motivational considerations in the study of instruction. *Cognition and instruction*, 5(4):289–309, 1988.

[55] Patrick Lin. Why ethics matters for autonomous cars. In *Autonomous Driving*, pages 69–85. Springer, 2016.

[56] Patrick Lin, Keith Abney, and George A Bekey. *Robot ethics: the ethical and social implications of robotics*. MIT press, 2011.

[57] Linda S Lumsden. Student motivation to learn. eric digest, number 92. 1994.

[58] Joseph J Martocchio and Jane Webster. Effects of feedback and cognitive playfulness on performance in microcomputer software training. *Personnel Psychology*, 45(3):553–578, 1992.

[59] Sandra Mathison. Why triangulate? *Educational researcher*, 17(2):13–17, 1988.

[60] Judith L Meece and M Gail Jones. Gender differences in motivation and strategy use in science: Are girls rote learners? *Journal of Research in Science Teaching*, 33(4):393–406, 1996.

[61] Joan Meyers-Levy and Barbara Loken. Revisiting gender differences: What we know and what lies ahead. *Journal of Consumer Psychology*, 25(1):129–149, 2015.

[62] Joan Meyers-Levy and Durairaj Maheswaran. Exploring differences in males' and females' processing strategies. *Journal of Consumer Research*, 18(1):63–70, 1991.

[63] Omar Mubin, Catherine J Stevens, Suleman Shahid, Abdullah Al Mahmud, and Jian-Jie Dong. A review of the applicability of robots in education. *Journal of Technology in Education and Learning*, 1:209–0015, 2013.

[64] Choon Boon Ng, Yong Haur Tay, and Bok Min Goi. Vision-based human gender recognition: A survey. *arXiv preprint arXiv:1204.1611*, 2012.

[65] Jakob Nielsen. Why you only need to test with 5 users, https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/, March 19, 2000. Accessed: 2017-04-27.

[66] Tatsuya Nomura, Takayuki Kanda, and Tomohiro Suzuki. Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *Ai & Society*, 20(2):138–150, 2006.

[67] Frank Pajares and M David Miller. Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of educational psychology*, 86(2):193, 1994.

[68] Tuomas Pekkarinen. Gender differences in education. *Nordic Economic Policy Review*, 1:165–195, 2012.

[69] Jan L Plass, Bruce D Homer, and Charles K Kinzer. Playful learning: An integrated design framework. *White Paper*, 2, 2014.

[70] Aaron Powers, Adam DI Kramer, Shirlene Lim, Jean Kuo, Sau-lai Lee, and Sara Kiesler. Eliciting information from people with a gendered humanoid robot. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 158–163. IEEE, 2005.

[71] Best Reviews. Best roombas, http://bestreviews.com/best-roombas, June 21, 2017. Accessed: 2017-06-22.

[72] Karsten Rex and Roberta M Roth. The relationship of computer experience and computer self-efficacy to performance in introductory computer literacy courses. *Journal of research on computing in education*, 31(1):14–24, 1998.

[73] Matt Richardson and Shawn Wallace. *Getting started with raspberry PI.* " O'Reilly Media, Inc.", 2012.

[74] Per Runeson, Martin Host, Austen Rainer, and Bjorn Regnell. *Case study research in software engineering: Guidelines and examples.* John Wiley & Sons, 2012.

[75] Martin Saerbeck, Tom Schut, Christoph Bartneck, and Maddy D Janse. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1613–1622. ACM, 2010.

[76] Eleanor Sandry. *Robots and communication.* Springer, 2015.

[77] Paul Schermerhorn, Matthias Scheutz, and Charles R Crowell. Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 263–270. ACM, 2008.

[78] Jean Scholtz. Theory and evaluation of human robot interactions. In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, pages 10–pp. IEEE, 2003.

[79] Mikey Siegel, Cynthia Breazeal, and Michael I Norton. Persuasive robotics: The influence of robot gender on human behavior. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 2563–2568. IEEE, 2009.

[80] Chris MR Smerecnik, Ilse Mesters, Math JJM Candel, Hein De Vries, and Nanne K De Vries. Risk perception and information processing: The development and validation of a questionnaire to assess self-reported information processing. *Risk Analysis*, 32(1):54–66, 2012.

[81] Wall street journal. Alexa, siri, cortana: Why all your bots are female, http://www.wsj.com/video/alexa-siri-cortana-why-all-your-bots-are-female/9e6e58c8-7b09-4de0-a0a0-a39e822d941b.html, February 21, 2017. Accessed: 2017-06-21.

[82] Amos Tversky and Daniel Kahneman. Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232, 1973.

[83] Alice M Tybout and Patricia Cafferata. *Cognitive and affective responses to advertising*. Lexington Books, 1989.

[84] MARJA VAN DEN HEUVEL-PANHUIZEN. Girls and boys problems. *International perspectives on learning and teaching mathematics*, pages 237–52, 2004.

[85] Ch K Volos, N Doukas, IM Kyprianidis, IN Stouboulos, and TG Kostis. Chaotic autonomous mobile robot for military missions. In *Proceedings of the 17th International Conference on Communications*, 2013.

[86] Kazuyoshi Wada and Takanori Shibata. Living with seal robotsits sociopsychological and physiological influences on the elderly at a care house. *IEEE Transactions on Robotics*, 23(5):972–980, 2007.

[87] Anthony Weber and Andrew Jones. Amazon echo: 2 in 1. the best user guides to learn amazon echo (alexa kit, amazon prime, users guide, web services, digital media, free books, free. 2016.

[88] Elke U Weber, Ann-Renee Blais, and Nancy E Betz. A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of behavioral decision making*, 15(4):263–290, 2002.

[89] C Wohlin, P Runeson, M Host, MC Ohlsson, B Regnell, and A Wesslen. Experimentation in software engineering: an introduction. 2000, 2000.

[90] J Yang and JF Coughlin. In-vehicle technology for self-driving cars: Advantages and challenges for aging drivers. *International Journal of Automotive Technology*, 15(2):333–340, 2014.

[91] Muhammad Ali Yousuf. Robots in education. In *Encyclopedia of artificial intelligence*, pages 1383–1388. IGI Global, 2009.

[92] Wenhao Zhang, Melvyn L Smith, Lyndon N Smith, and Abdul Farooq. Gender and gaze gesture recognition for human-computer interaction. *Computer Vision and Image Understanding*, 149:32–50, 2016.

APPENDICES

Appendix A: Pre-task Session Questionnaire

Pre-task form

Please write down the following information:

Age

Gender

Degree major

Video gaming experience

Often in our jobs we are told to use hardware that are available to make our work easier. For the following questions, imagine that you were given a robot for some aspect of your work. It doesn't matter specifically what this robot does, only that it is intended to make your job easier and that you have never used it before. The following questions ask you to indicate whether you could use this unfamiliar robot under a variety of conditions. For each of the conditions, please indicate whether you think you would be able to complete the job using the robot. Then, for each condition that you answered "yes", please rate your confidence about your first judgment, by circling a number from 1 to 10, where 1 indicates "Not at all confident", 5 indicates "Moderately confident", and 10 indicates "Totally confident".

I could complete the task using robot ...

No.	Question	Not at all confident				Moderately confident				Totally confident	
Q-1	...if there was no one around to tell me what to do as I go	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-2	...if I had never use a robot like this before	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-3	...if I had only the robot manual for reference	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-4	...if I had seen someone else using it before trying it myself	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-5	...if I could call someone for help if I got stuck	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-6	...if someone else had helped me get started	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-7	...if I had a lot of time to complete the task for which the robot was provided	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-8	...if I had just the built in help facility to get started	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-9	...if someone showed me how to do it first	YES…1 NO	2	3	4	5	6	7	8	9	10
Q-10	...if I had used similar robot before this one to do the same task	YES…1 NO	2	3	4	5	6	7	8	9	10

Imagine that you are asked ***to move a delicate expensive flower vase*** on the table using the PR2 robot. Now write down answers to the following questions for moving the flower vase using robot.

Q1. What information do you need to complete the task using robot?

Q2. Suppose you are provided with a handout and a video on how to use the robot. Which option would you choose? And why?

Q3. How are you going to use the information provided in the handout and the video?

Q4. Suppose you are given a good sum of money, would you buy this robot? Mention why/why not?

Q5. How often do you play with a new hardware device? Why? Why not?

Q6. How likely are you going to experiment with the robot to complete the given task?

Appendix B: Post-task Session Questionnaire

Post-task form

Now that you have used the robot to complete given tasks please answer the following questions. Your answer should be based on your experience with the robot and the materials (handout and video) provided on how to use the robot. For each condition that you answered "yes", please rate about your first judgment, by circling a number from 1 to 7, where 1 indicates "Slightly", 4 indicates "Moderately", and 7 indicates "Absolutely".

No.	Items	Slightly			Moderately			Absolutely
Q-1	I thought about what actions I might take based on what I read and/or saw on the video	YES… 1 NO	2	3	4	5	6	7
Q-2	I found myself making connections between the information and what I've read or heard about elsewhere	YES… 1 NO	2	3	4	5	6	7
Q-3	I thought about how the information related to other things I know	YES… 1 NO	2	3	4	5	6	7
Q-4	I tried to think about the importance of the information	YES… 1 NO	2	3	4	5	6	7
Q-5	I tried to relate the ideas in the information	YES… 1 NO	2	3	4	5	6	7
Q-6	I skimmed through the information	YES… 1 NO	2	3	4	5	6	7
Q-7	I did not spend much time thinking about the information	YES… 1 NO	2	3	4	5	6	7
Q-8	The given handout and/or video did not contain useful information on which I based my decision	YES… 1 NO	2	3	4	5	6	7
Q-9	While reading the information, I did not think about the advices presented in the information	YES… 1 NO	2	3	4	5	6	7
Q-10	The information contained too many incompatible viewpoints	YES… 1 NO	2	3	4	5	6	7

Answer the following questions based on your experience with the ***final task***.

Q1. What information did you use to complete the task?

Q2. How did you use that information?

Please rate on a scale of 1 to 10.

How confident you were that you had successfully moved the …

No.	Question	Not at all confident				Moderately confident					Totally confident
Q-1	… robot arm?	☐ 1	2	3	4	☐ 5	6	7	8	9	☐ 10
Q-2	… robot hand in a proper orientation?	1	2	3	4	5	6	7	8	9	10
Q-3	… robot fingers at proper position?	1	2	3	4	5	6	7	8	9	10

No.	Question	Very bad				Moderately good					Very good
Q-1	How would you evaluate your performance?	☐ 1	2	3	4	☐ 5	6	7	8	9	☐ 10

No.	Question	Not at all				Moderately					Very much
Q-1	How often did you make any risky move that you were not sure of?	☐ 1	2	3	4	☐ 5	6	7	8	9	☐ 10
Q-2	How did you enjoy the task using the robot?	1	2	3	4	5	6	7	8	9	10

Appendix C: Handout Tutorial

1. **You will be using the following robot shown in Figure C.1. The robot is ready to perform tasks.**

2. **To perform given tasks, you will need to know only the following and feel free to ask any questions that you might have.**

    - To perform the tasks (moving several objects), you have to move robot arm, hand, and fingers. Heres the reference point device shown in Figure C.2, you will have to make sure you are standing beside this device to control the robot.

    - While performing the task please be cautious and make sure that the robot hands doesnt hit each other or with anything else.

    - You can also ask for a video where you will see how PR2 is operated by a human. And/or choose to use this handout.

    - You have to use one/both joysticks (for left and right hand) shown in Figure C.3 and instruction is provided in Figure C.4. Please make sure you understand the functions of each buttons (only marked ones) of the joysticks for correct operation.

Figure C.1: PR2 robot with its hand in a certain pose.

Figure C.2: Reference point device working (green color light on).

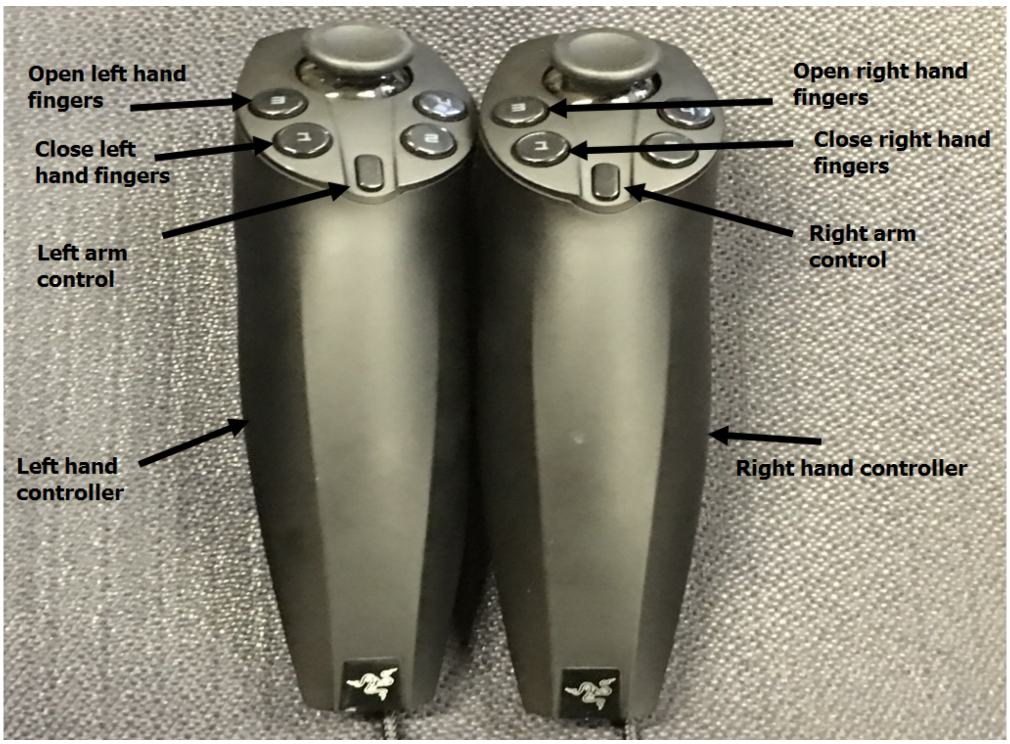

Figure C.3: Joystick for moving robot arm, hand and fingers.

Button	**Control Action**
1 (both for left and right arm)	Pressing 1 on the left joystick will **close** the left-hand fingers, pressing 1 button on the right joystick will **close** the right-hand fingers.
3 (both for left and right arm)	Pressing 3 on the left joystick will **open** the left-hand fingers, pressing 1 button on the right joystick will **open** the right-hand fingers.
Center button (left)	Switch for controlling the left arm. While pressed, the left arm will mimic the movements of the left hand. (keep this button pressed so that it mimics your left arm movements)
Center button (right)	Switch for controlling the right arm. While pressed, the right arm will mimic the movements of the right hand. (keep this button pressed so that it mimics your right arm movements)

Figure C.4: Joystick control description.

Appendix D: Data and Coding

Tinkering open-ended question response and coding:

Gender	*Quotes*	*Code*
Male	"I love trying new devices such as VR, tablets, smart watches."	pf
Male	"Often, because it is fun."	pf
Female	"I dont play with new hardware device, I am a peoples person."	np
Female	"Almost never. I am not tech savvy at all."	np
Female	"I dont have a lot of money."	np
Male	"I love to play with all new gadgets come up in the market."	pf
Male	"I Try out new device once a month, use ardunio ."	pf
Male	"I love to see how technology is developing and understand how to use it."	pf
Male	"Its fun and interesting to use."	pf
Female	"Not much."	np
Female	"Not often."	np
Female	"Not often, they are expensive."	np

Table D.1: Response from tinkering open ended question and coding (playful = pf, not-playful = np).

Codes	*Description*	*Object*
Top-left/right	Pick up object from top using left/right hand	Wine glass
Stem-left/right	Pick up object from bottom using left/right hand	Wine glass
Tip-left/right	Pick up objects tip using left/right hand	Plastic bowl
Body-left/right	Pick up object body with left/right hand	Wine glass
Claw-horizontal-left/right	Left/right claw placed horizontally	Rose, scarf
Claw-vertical-left/right	Left/right claw is placed vertically	Rose, scarf
Claw-both	When using both the hands	Scarf

Table D.2: Tinkering codes with object annotation.

Codes used for tinkering from video:

Gender	*Bowl*	*Rose*	*Wine Glass*	*Scarf*	*Total Tinkering Count*
Male	1	2	3	2	8
Female	1	1	2	1	5
Male	2	1	2	3	8
Female	1	1	1	2	5
Female	1	1	1	1	4
Male	1	3	1	6	11
Female	2	1	1	2	6
Male	2	1	1	3	7
Male	2	1	1	1	5
Male	1	1	2	3	7
Female	1	1	1	2	5
Female	1	1	1	2	5

Table D.3: Participant's tinkering count summary while performing tasks using robot to manipulate various objects.

Coding summary of tinkering with robot from video data:

Gender	*Questions Asked*
Male	N/A
Female	"how to turn it on? How to control it?"
Male	"how to move the robot? How to access handling parts?"
Female	"how the robot moves?"
Female	"How I give it commands?"
Male	"How do I move the arm? How do I open the claw? How do I rotate the claw? "
Female	"How it works? "
Male	"how to move the objects? How to do so slowly?"
Male	"How to hold the vase?"
Male	N/A
Female	"How to start the robot?"
Female	"How to respond to it and how it responds back?"

Table D.4: Strategy type question asked by participants as response to open ended question..

List of Strategy questions asked for using the robot:

Gender	*Questions Asked*
Male	N/A
Female	“how fast it moves? Which directions the arms move?”
Male	N/A
Female	“how far the robot may have to travel to be in front of the object? ”
Female	“what it uses to grab stuff?”
Male	N/A
Female	N/A
Male	N/A
Male	“how fragile the vase is?”
Male	N/A
Female	“what keys to use? Which button do each task?”
Female	“how big the object is so to accommodate the robots hand to pick it up?”

Table D.5: Feature type question asked by participant as response to open ended question.

List of Feature type questions asked for using the robot:

Gender	*Questions Asked*
Male	"How fast it rotates?"
Female	"how its hands function?"
Male	N/A
Female	"how capable it is to pick things correctly?"
Female	"How it moves across the floor? How it put stuff down?"
Male	N/A
Female	"How accurate are the controls are?"
Male	N/A
Male	N/A
Male	If there's overshot, where to stop it?
Female	"How sensitive the controls are?"
Female	"Whether or not robot would have to rotate itself?"

Table D.6: Functionality type question asked by participant as response to open ended question.

List of Functionality type questions asked for using the robot:

Gender	*Bowl*	*Rose*	*Wine Glass*	*Scarf*	*Total*
Male	100	0	100	80	280
Female	100	100	100	20	220
Male	100	100	80	80	360
Female	100	100	20	0	220
Female	100	100	0	0	200
Male	100	100	100	80	380
Female	100	100	0	20	220
Male	100	100	100	100	400
Male	100	100	100	100	400
Male	100	100	0	80	280
Female	0	100	0	20	120
Female	100	100	0	80	280

Table D.7: Task score summary while performing tasks using robot to manipulate various objects.

Coding summary of task score from video data:

Gender	*Bowl (sec)*	*Rose (sec)*	*Wine Glass (sec)*	*Scarf (sec)*	*Total (sec)*	*Total (min)*
Male	97	201	234	197	918	15.30
Female	115	264	508	176	1063	17.71
Male	202	134	131	148	615	10.25
Female	92	243	334	806	1475	24.58
Female	170	403	153	689	1415	23.58
Male	92	144	33	833	1102	19.36
Female	383	213	312	557	1465	24.41
Male	130	77	75	178	460	7.66
Male	154	269	86	121	630	10.5
Male	159	136	153	309	757	12.61
Female	40	576	30	741	1387	23.11
Female	91	257	220	388	957	15.95

Table D.8: Task time summary while performing tasks using robot to manipulate various objects.

Coding summary of task time from video data: