Sociable Robots: The Role of Presence and Task in Human-Robot Interaction

by

Cory David Kidd
Bachelor of Science, Computer Science
Georgia Institute of Technology, May 2000

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2003

Signature of author: ___
Cory D. Kidd
Program in Media Arts and Sciences
May 2003

Certified by: ___
Dr. Cynthia Breazeal
LG Group Career Development Professor
Assistant Professor of Media Arts and Sciences
MIT Media Laboratory

Accepted by: ___
Dr. Andrew Lippman
Chair, Departmental Committee on Graduate Students
Program in Media Arts and Sciences

Sociable Robots: The Role of Presence and Task in Human-Robot Interaction

by

Cory David Kidd

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning on May 9, 2003
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

Human-robot interaction is an emerging area of study that is developing an understanding of how to build robots that are useful and effective in helping people perform tasks in particular domains. In particular, social robots, or those that help people as capable partners rather than as tools, are believed to be of greatest use for applications in entertainment, education, and healthcare because of their potential to be perceived as trusting, helpful, reliable, and engaging. This thesis explores how the robot's physical presence and proximity to a person influence a person's perception of these characteristics. Results from two experiments are reported: the first shows differences in participant responses to a robot, an animated character, and a human and the second shows the outcome of participants interacting with a robot or a robot presented on a television in two different types of tasks. Responses to the interactions were collected via questionnaire and videotape and are reported on scales measuring trust, perceived information quality, altruism, level of engagement, reliability, immediacy, credibility, and persuasiveness. The results of this research will contribute towards the goal of building robots that can effectively communicate with and assist humans in a variety of applications in domains that I believe will benefit from social robots.

This work was sponsored in part by a Graduate Research Fellowship from the National Science Foundation

Thesis Supervisor: Cynthia Breazeal
Title: Assistant Professor of Media Arts and Sciences

Sociable Robots: The Role of Presence and Task in Human-Robot Interaction

by

Cory David Kidd

Thesis Advisor: ______________________________
Dr. Cynthia Breazeal
LG Group Career Development Professor
Assistant Professor of Media Arts and Sciences
MIT Media Laboratory

Thesis Reader: ______________________________
Dr. Rosalind Picard
Associate Professor of Media Arts and Sciences
MIT Media Laboratory

Thesis Reader: ______________________________
Dr. Clifford Nass
Professor
Stanford University Department of Communication

Acknowledgements

First, many thanks to my advisor, Cynthia Breazeal, for getting me started on the path that this research has taken. Her inspiration through two years of conversations on what robots could do have been behind the work that is presented here. I am looking forward to several more years of exploration and research in human-robot interaction under her guidance.

The guidance of two other people has been invaluable in the conception and design of my experiments: Roz Picard of the Media Lab and Cliff Nass of Stanford University. Appropriately, both of them have become readers on this thesis. Thank you to both of them for their many ideas over the last year and a half.

I would also like to thank the rest of the Robotic Life group, both past and present: Andrew "Zoz" Brooks, Matt Hancher, Jeff Lieberman, John McBean, Dan Stiehl, and Josh Strickon. You have all made this research group a fun and interesting one to be a part of for the last two years.

Thanks also to the UROPs who have worked with me on assorted research endeavors over the last couple years: Ming-Fai Fong, Sam Hwang, Heather Knight, Adam Kraft, and Tuan Phan. And especially to Michael Wolf, who has spent many, many hours collecting the data from the two human-robot interaction experiments. I could not have completed the amount of analysis I did without his help.

Additionally, the work of Ryan Kavanaugh, Geoff Beatty, Marc Downie, and Matt Berlin in creating the animated character for my first experiment was extremely helpful. Thanks to them and the rest of the Synthetic Characters group for their contribution to this thesis.

In the Media Lab, I would also like to thank Linda Peterson and Pat Solakoff for seeing that I got through these two years and even got paid most months. Don't worry, I'll be around for a few more years to see that we've got it all worked out.

There are a few other people around the lab who I would like to thank for their conversations, support, and general friendship: Lily Shirvanee, T.J. McLeish, Michael Rosenblatt, and Liz Sylvan. I'm glad that you will all be at the lab or at least in the Boston area in the coming years.

Finally, I would like to thank my family for encouraging me through these many years of school. My fiancée, Erica Young, has been a tremendous source of support and encouragement throughout all of this work. Not to mention her serving as the voice of my robots through two experiments. I appreciate her interest in my research and her willingness to put up with many late nights and working weekends. I promise, just one more degree and I'm done – for now!

Table of Contents

1 Introduction 13

1.1 Contents of This Thesis 13
1.2 Robots as Partners 14
1.2.1 Current and developing robotic domains 15
1.2.2 Future robotic domains 17
1.3 Types of Social Robots 24
1.4 What Will It Take? 25
1.5 Important Design Issues 25

2 Background 27

2.1 Humans and Computers: HCI and HRI 27
2.1.1 *The Media Equation* – inherent social responses to media 27
2.1.2 Presence 28
2.1.3 Mental models of robots 29
2.1.4 Robot personality 29
2.1.5 Appearance of robot 30
2.1.6 Gender effects 30
2.1.7 Credibility 31
2.1.8 Liking and Preference 31
2.1.9 Perceived information quality 31
2.2 Human Behavior: Social Psychology and Sociology 32
2.2.1 Presence 32
2.2.2 Sociability 34
2.2.3 Social judgments 35
2.3 What Does This Mean for Robots? 35
2.4 Overview of Work Presented Here 37

3 First Experiment 39

3.1 Relevant Measures 39
3.1.1 Arousal 39
3.1.2 Social Presence of Characters 40
3.2 Experimental Methods 40
3.2.1 Participants 40
3.2.2 Experimental Design 41
3.2.3 Protocol 48
3.2.4 Dependent measures 49
3.3 Results 50
3.3.1 Qualities of the Interactions 50
3.3.2 Participant Responses to Interactions 53
3.3.3 Engagement with Characters of Different Media 54
3.3.4 Description of Characters 55
3.3.5 Social Presence 57
3.3.6 Electrodermal Response 60

3.3.7 Video Tape Data 61
3.3.8 Robot as Compared to Other Modalities 61
3.4 Summary 62

4 Second Experiment 63

4.1 Relevant Measures 64
4.1.1 Trust 64
4.1.2 Perceived Information Quality 64
4.1.3 Altruism 65
4.1.4 Engagement 65
4.1.5 Reliability 66
4.1.6 Immediacy 66
4.1.7 Credibility 66
4.1.8 Persuasiveness 67
4.2 Experimental Methods 67
4.2.1 Participants 67
4.2.2 Experimental Design 68
4.2.3 Protocol 74
4.2.4 Dependent measures 79
4.3 Results 79
4.3.1 Overview 80
4.3.2 Task 81
4.3.3 Presence 82
4.3.4 Gender 85
4.3.5 Interactions Among Task, Presence, and Gender 86
4.3.6 Distance from robot 87
4.3.7 Time looking at robot 87
4.3.8 Time of Interactions 88
4.3.9 Teaching task outcome 88
4.4 Summary 88

5 Discussion 91

5.1 Experimental Questions 91
5.2 Effects of Modality 92
5.3 Effects of Proximity 93
5.4 Effects of Interactivity and Task Type 94
5.4.1 Task outcomes 94
5.5 Effects of Gender 96
5.6 What Does This Mean? 97
5.6.1 Agreement with Previous Research 98
5.7 Summary 100

6 Conclusion 101

6.1 Applicability to Particular Domains 101
6.2 Robots as Partners 102

6.3 Future Work 102

References 105

Appendix A – Experiment I Questionnaire 111

Appendix B – Experiment II Questionnaire 121

Subscale: Trust 121
Subscale: Perceived Information Quality 121
Subscale: Altruism 122
Subscale: Engagement 122
Subscale: Reliability 123
Subscale: Immediacy 123
Subscale: Credibility 124

Appendix C – Teaching Task 127

Appendix D – Experiment II Data 131

Measure: Trust 131
Measure: Altruism 131
Measure: Engagement 132
Measure: Reliability 133
Measure: Credibility 134
Measure: Perceived Information Quality 135
Measure: Immediacy 135
Measure: Distance 136

1 Introduction

Robots will soon be a part of our everyday lives. The field of robotics research is advancing to the point where it is becoming easier to build robots that can interact with us in our offices, homes, schools, and research labs. Robots will soon fill more roles than their current applications as industrial assemblers and children's toys. We have now reached a point where we are able to begin building robots that can act as partners with us: teachers, healthcare and domestic assistants, actors, scientific collaborators, and entertaining playmates.

What does it mean to have a robot act as a partner instead of simply a tool or a toy? In general, it means that we can expect the robot to act and react in many ways that a human can: understand our directions to complete a task, guide us in learning something new, and assist us when we need a helping hand. The following five sections highlight two types of these applications, both current work in robotics that is already trying to make these robots a reality and ideas for the future that we expect to see soon.

1.1 Contents of This Thesis

This first chapter gives the motivation behind the work presented in the remainder of the thesis and highlights the key points that will be discussed in coming chapters.

The second chapter presents the background work that is related to this thesis and will give a grounding for the analysis that will be presented in later chapters. If you want to understand how the work presented in this document fits in to the larger body of research in human-computer interaction and, more specifically, human-robot interaction, the second chapter will give you that information.

The following two chapters are for those who really want to understand the specifics of the work that I have carried out over the past two years. These two chapters are on the experiments that were conducted, one chapter discussing each of the experiments in detail. I present a complete discussion of the design rationale, the implementation and protocol, and a summary of results.

Finally, we get to the most interesting chapters of this thesis from a theoretical perspective. The Discussion chapter points out the meaningful and significant results from the two experiments and talks about how they impact the field of human-robot interaction. There is less data presented here than in the experiment chapters and much more relating these results to real-world examples.

The Conclusion gives a concise summary of the analysis presented in the previous chapter and ties it to the larger body of work.

1.2 Robots as Partners

Two models have been proposed for how we can develop and use the robots that we are working to build. One way is to consider the robot as a tool: it is an object that is meant to be manipulated by a human. In this case, the robot is conceptually similar to a screwdriver or a wrench. A person can directly use the robot in some way to manipulate objects in the world. With complicated robots, these manipulations may be complex, such as using Intuitive Surgical's da Vinci robotic surgical system [46]. Regardless of the complexity, however, we do not expect these robots to do much on their own without the input of their human controller.

Contrast this with the concept of the robot as a partner in accomplishing a task. In this scenario, the robot takes direction from a human at a high level and carries out the low-level operations that are necessary to complete a task. To illustrate this idea, take one of the examples that will be discussed in greater detail shortly, a robot as an astronaut. In this case, a robot might be directed to go outside a space vehicle to attach a new piece of equipment to the International Space Station. The robot would then know how to find the location, attach and test the piece, and return to the space vehicle without direct human intervention at each step. For a closer-to-home example, imagine a robot in your home that acts as a personal assistant. If you had to direct the robot at every step of the way to take out the garbage each time it was to do so, the robot would not be seen as helpful by very many people. In contrast, if you can request that the robot take out the trash and it could find the garbage can, remove the bag, and place it in the right location on the curb, that would be seen as very helpful to have around your home.

1.2.1 Current and developing robotic domains

The most notable current use of robots is in entertainment. Most people have seen or are somewhat familiar with Sony's AIBO, Hasbro's My Real Baby, or the numerous other robotic toys for children. Another use of robots in entertainment is as a robotic actor in a movie. Although these applications appear on the surface to be vastly different than some of those discussed in the coming pages, such as a robotic astronaut or a healthcare assistant, there are actually a number of similarities across these applications in how the robot must be able to comprehend and react to a human with whom it is interacting. The area of research concerned with these interactions is termed social robotics and human-robot interaction. A good overview of many of the aspects of human-robot interaction and the uses of social robots can be found in a recent review by Fong, Nourbakhsh, and Dautenhahn [20].

Entertainment

The first place where many of us have already seen robots used is in the entertainment industry. The commercial success of creature-like robots such as the Sony Aibo dog and the multitude of less expensive (and less-abled) toys that the Aibo inspired suggests that the ability to make a toy with true social interaction capabilities could create an even more compelling and better-selling toy. If we look at just one example of a current robotic toy, Hasbro's My Real Baby, we can see how quickly many people become attached to this toy. The work of Sherry Turkle and Olivia Dasté [24, 47] has shown a number of examples of both children and adults developing an attachment with one of these robotic dolls. While this emotional attachment to a robot may not always be desirable, it may promote the robot's effectiveness in certain situations, such as a robot designed to assist a child in her school lessons or one that is meant to help an autistic child adapt to life in the real world [45]. This willingness to patiently teach and assist a robot will clearly be an asset when interacting with the first generations of consumer-focused robots that will inevitably need a great deal of human hand-holding to learn how to perform its tasks around the home or office. This patience will also be a virtue when robots make mistakes – not many of us feel a lot of patience for our computers, which leads to frustration and even a desire to hit or kick them on occasion [40].

There is also a strong interest from the Hollywood entertainment industry. There is already a great use of robots on the silver screen (for example, the *Jurassic Park* series and the *Terminator* movies). The challenge in directing the current generation of robots on a film set comes from the complexity of their design. Most robots are composed of dozens of individual motors (or degree of freedom, abbreviated DOF) which must be controlled in real-time by several people, usually the engineers that built the robot. Because each person is limited in how many DOFs they can control (typically 8-10 per person), this leads to the need for a team of people to control an individual robot. This set of issues causes two problems that are inherent in this kind of setup. The first is that it is challenging for even an experienced controller to get realistic movement out of a robot. Designers have discovered that it takes on the order of ten DOFs around the mouth to get movements that are believable-looking as speech. This means that one person must control all of these DOFs at once, which is not a simple task. A controller spends weeks, if not months, perfecting his technique to be able to produce realistic-looking speech.

Apart from the complexity of one person controlling movements to make their part of the robot appear realistic comes the challenge of choreographing up to eight people controlling a single robotic character's performance at one time. Among the people who control these robots on the set, it is known as a challenging problem to do things well that we find trivial for ourselves, such as keeping a character's eyes fixed on a point (another actor to whom he is speaking, for example) while turning his head. This is because typically a different person will be controlling the eyes than the person who is operating the overall head and neck. If the robot had the ability to perform this kind of skill on its own, the complexity of the control problem would be reduced and the performance could be closer to the ideal of the single-minded human actor.

A robotic character on a set also provides an advantage to the actors who must perform opposite them when the alternative is a blue- or green-screen performance. (In this situation, the actor performs in front of a solid blue or green wall and the other "character" is digitally added into the scene in postproduction [50].) Actors, and indeed all of us, typically find it easier to respond to someone or something that is in front of us instead of only in our minds. The same argument can be made for the character knowing how to perform natural social responses such as orienting to the person who is talking to it or to be able to share visual attention cues with their co-star.

1.2.2 Future robotic domains

The more compelling and interesting applications are further away and somewhat more challenging to achieve. The science-fiction-inspired vision of the personal robot for performing our mundane household chores will be a difficult one to achieve without the advances sought after through sociable robotics. Imagine trying to explain to another person how you clean your kitchen or fold your laundry using a traditional graphical interface. Even if it's not impossible, it hardly seems desirable.

There are other approaches to this problem that have been suggested, such as a system that simply learns by observation without providing feedback. We believe, however, that the robot's ability to understand and display social interaction cues with a person is an important feature that will allow this kind of robot to be more successful in conducting interactions. If a person feels that the robot is trying to communicate with them and learn from them, they will be more inclined to help it succeed.

Human-robot teams for scientific exploration

A near-term project of this type of robotic assistant is the Robonaut project that is currently in progress [34]. Researchers at the National Aeronautics and Space Administration in the United States (NASA) are currently working with a number of other research groups to build a robot that will assist astronauts with various tasks in space. The purpose of this robotic astronaut is "to develop and demonstrate a robotic system that can function as an EVA [extra-vehicular activity, or space walk] astronaut equivalent." [35]

Another key design feature of this robot (and for social robotics in general) is for the robot to be able to learn through natural interaction with a human. The Robonaut proposal gives three important reasons why these capabilities will make the system easier to use and more successful. The first concerns the cognitive load that would be placed on an operator trying to control a 47 degree-of-freedom robot. This is an immensely complex task that requires the operator to wear a virtual reality immersion helmet for visual feedback and sensing the position and orientation of the operators' head, gloves for sensing the position of each finger, and an exoskeleton for determining the positions of other body segments. The visual feedback received by the operator through the display in the helmet is the only feedback that is received, so the operator must be very attentive to what is being displayed. The combination of the equipment that is required to be worn and the constant, careful attention needed of the operator means that one person can only effectively control the robot for a short period of time if it is to be teleoperated.

The second reason that social interaction capabilities are desired is that training the robot will be simpler from the human teacher's perspective. Similar to the problem of controlling each of the degrees of freedom on this robot, programming the robot can be equally challenging. A particular type of social interaction that would be useful here is learning by imitation or other forms of social learning. If the robot is capable of this type of learning, rather than having each joint programmed for each activity that the robot is to perform, it can watch a human perform the task and then imitate that action. The human trainer can then correct aspects of the robot's action if necessary

The final reason for implementing social interaction in this robot is that "the social aspects [of] human pedagogy will allow the robot to learn more effectively from fewer examples and generalize better." This aspect of learning takes the example of imitation learning a step further. Instead of learning particular actions and then being told when and where to carry out those actions, the robot could learn higher-level goals concerning the type of work that it is supposed to conduct. For example, the robot could learn that it should look for anomalies in particular systems and then correct them whenever they occur.

Both the second and third issues are discussed further in Breazeal's 2002 book on designing sociable robots [10]. Clearly this kind of robot could provide great benefit to human endeavors in space and will best be achieved through the application of the principles of human-robot interaction.

Human-robot teams for search and rescue

Another currently developing use of robotics is in search and rescue missions. In these scenarios, we can think of a robot in much the same way as we would a dog. Dogs are commonly used as search and rescue teammates because of their intelligence and ability to be trained. We want two things from these dogs: they should be able to follow commands and general guidelines that let them know where they should go and what kind of things they should do (search for people and alert human rescuers, for example). They must also have the intelligence to carry out this search on their own without constant direction by a human, moving completely throughout a space that may be too hazardous or too small for a human to search [16].

This application of robots blurs the distinction between viewing the robot as a tool or as a partner. In some sense, it is simply a tool that we are using to complete a particular task that may not be desirable or even possible for a human to accomplish alone. However, the partnering aspect of this task – where the robot takes orders and autonomously carries out a search – is clearly an important aspect contributing to the success of these robots. In an urgent, highly stressful situation such as search and rescue, it is vital that these robots be able to quickly, easily, and reliably interact with humans in as natural a manner as possible so that the desired outcome is achieved.

Household robots

Although there are no existing robots that fall into the category of social household robots, there have been attempts to create commercially viable, but simpler, robots that assist with household tasks. The most recent of these is the Roomba vacuum cleaner robot from iRobot [22]. According to the iRobot web site, the Roomba "is the first automatic vacuum in the U.S. It uses intelligent navigation technology to automatically clean nearly all household floor surfaces without human direction." While this robot does not demonstrate the viability of the ideas presented in this thesis, its commercial success does show that there is a growing acceptance of robots as a household assistant.

With the capabilities that social interaction can provide, it will be possible to create robots that perform more complex tasks around the house than vacuuming the carpet. One of the limiting factors of the current generation of vacuuming robots is that they have difficulty knowing where to and where not to vacuum. This is because the input to them is usually limited to a switch that turns them on and off. It's then up to the robot to figure out where to go based on algorithms that must be pre-programmed when the robot is designed and constructed. If the robot had the capability to quickly learn where and when to do its job, the resulting product would be much more satisfying to the consumer.

Informational robots

One type of robot that is being studied but does not yet have a good name is robots that interact with people in public spaces and convey information. One example of this could be a robot that greets visitors in an office building or lab and gives directions or other information. Another example that has been implemented by a team of researchers at Carnegie Mellon University is a robot that guides visitors around a museum and tells them about the exhibits [20]. This robot, called Sage, wanders around the exhibits in Dinosaur Hall at the Carnegie Museum, telling visitors about the exhibits that they are near and helps direct them to other displays.

Social interaction is clearly important here because these robots will be in situations and locations where they will be interacting with many different people. In a controlled lab setting, users can be trained to interact with a robot in a way that the robot can interpret. In a public setting, the robot must understand how to interact with people in a very social manner. This includes things such as moving through crowds (something that the aforementioned robot, Sage, does), conversing at appropriate times, and understanding and conveying information that is appropriate to its audience.

Communicative robots

Many technologies have been developed to enable, support, and extend communication capabilities between people. The telegraph, telephone, television, and e-mail are all examples of these kinds of technologies. Each has its benefits, but there are also limitations to each of these means of communication. If we concentrate on two-way interaction between individuals, two of the most widely used of these technologies are currently the telephone and e-mail. The greatest difficulty encountered in interactions across either of these media is the lack of non-verbal channels of communication. There is extensive research showing that these channels (such as facial expression, body posture and movement, and eye gaze) are extremely important to engendering trust, liking, and other factors that are greatly important in any social task and important, albeit to a lesser extent, in non-social interactions [2, 7, 13, 15, 29]. Another difficulty is an obvious one: the inability to physically share the same space and manipulate objects or even point at some object between the participants in an interaction.

For many years (over 75 years, in fact) people have thought of using videoconferencing as a solution to these problems. The belief has been that the ability to see the other person (or people) involved in a conversation would open up these other, non-verbal channels of communication for use in a conversation. However, there are a number of shortcomings to this potential solution, many of which are discussed by Hannes Vilhjálmsson in his recent dissertation [49]. He points out that turn-taking is difficult because it relies on gaze direction, which is generally not supported in video communication systems; assessing a conversation partner's focus of attention is also challenging; side conversations cannot take place in groups because everyone is sharing the same communication channels; and pointing or manipulating physical objects is difficult. There are obviously difficulties that must be overcome in order to create an acceptable technology for communication at a distance.

One means for achieving this type of communication may be through robotic avatars. These avatars would serve as the remote embodiment of a person participating in an interaction. There are several advantages to this system over a videoconferencing setup. Gaze direction would be much simpler to interpret when there is a physically present character representing the other person (or people) in a conversation. (This is shown in the first experiment presented in this thesis.) This would help to alleviate the first two problems mentioned above: turn-taking regulation and determining focus of attention. Pointing and manipulating objects may also be possible depending upon the construction of the robot.

Control of these robots would of course be an issue. However, if we take advantage of some of the same ideas presented in the previous sections on robots for scientific collaboration or search and rescue, it becomes a simpler problem. We can imagine utilizing technology to allow the robot to synchronize its facial and gestural movements with the speech of the person it is representing. Manipulation of objects could be achieved through currently available means of telepresent operation or through similar mechanisms to those described in the scientific collaboration section, allowing a person to provide the high-level direction for the robot and leaving it to the robot to manage the detailed movement.

This kind of interaction certainly holds promise for the future of remote communication and collaboration. There remains a great deal of work to achieve these goals, but more research should be focused in this direction to understand not only what is possible, but what is desirable and beneficial in using robots for this kind of mediated communication.

Educational robots

Another very important application of sociable robotics is in education. There are currently a plethora of computer-based tutorials for students on a wide range of subjects. An important aspect of the mentor-student relationship is the shared reference through cues such as directing attention, mutual gaze, pointing, and displaying and reading facial expressions, features that computer-based tutorial systems do not currently possess. These social aspects of the mentor-student relationship are an important part of the learning process, so understanding how to create these as a part of an interaction with a robot is an important step towards creating robots that will successfully fill this kind of a role. When it is not possible to have a human mentor, or when the human mentor is at a distance (such as in remote learning scenarios), a robot may prove to be more engaging and easier to interact with than a computer-based tutor because of the shared physical space.

Healthcare robots

No less important than employing robots in education is their potential use in health care. As the population of the world is aging [48], the number of elderly needing regular attention is growing. With the shrinking number of working-age caregivers available, robots are a logical alternative for providing some portion of this care. There are a number of projects that seek to address this problem [4, 14], and our work contributes to an understanding of the characteristics that these robots should possess to make the interactions rewarding, or at least palatable, to their patients. A key feature of a robot in this domain is the ability for the person needing care to maintain a feeling of independence and not feel as though they are a burden to a caretaker; rather, they can have a robot act as an extension of themselves or as an assistant to aid them in their everyday tasks. It is also important to have a robot that a human feels they can trust, is useful, and has their best interests in mind.

A different aspect of health care related to robotics is the use of robots in pet therapy. The idea behind pet therapy is that people are happier, healthier, and recover faster from ailments when they have the company of a pet [1, 41]. One problem with a typical situation in which a pet might be beneficial is that the person who might benefit from the pet has difficulty taking care of a pet, either because of difficulties keeping a regular schedule, because of health reasons (allergies, etc.), or because their care environment does not allow pets (hospital or nursing home, for example). In these cases, it may be beneficial to have a robotic pet that could be cared for by the person. It may still be possible to make the emotional attachment that is desirable in this kind of relationship without some of the detriments that come with a living pet, especially in a clinical environment. There are already examples of companies trying to create and market this kind of robotic pet. A leading company in this respect is the Japanese company Omron with its NeCoRo pet [17]. The development of improved interaction capabilities could make this kind of robot more beneficial to its intended recipient and allow those who cannot have pets to attain the benefits that doing so would give.

1.3 Types of Social Robots

Cynthia Breazeal has delineated several types of social robots in her recent paper on the topic [9]. She charts the range of robots as they increase in their ability to interact with humans in more natural ways for the human – through more complex understanding and exhibition of social cues that humans normally use in communication.

This starts with *socially evocative* robots, robots that are created in such a manner as to encourage people to treat them as animate objects. Examples of this kind of robot are toys that are designed to draw people in to an interaction, but do not go further in engendering a social alliance.

Next come robots described as having a *social interface*. These are robots that can exhibit social cues that a human would recognize. Alongside this category are *socially receptive* robots. These are robots that can comprehend social cues that are displayed by people with whom the robot is interacting. An important feature of these robots according to Breazeal's definition is that they should be capable of learning from this social interaction and benefit from level of understanding.

Finally, we have the category that we are most concerned with in the current work, that of *sociable robots*. These are robots that are capable of fully engaging in social situations and have motivations of their own for being involved in these types of interactions. This type of robot will be capable of embodying all of the traits described for all of the kinds of robots described above. This type of robot will be capable of interacting on a social level that humans are familiar and comfortable with, thus allowing these interactions to take place naturally and easily.

1.4 What Will It Take?

It is obvious that the type of robot discussed here does not yet exist. The type of interaction that is often touted as the gold standard for human-robot interaction comes from the set of principles that are known about humans interacting. This is the type of study that usually falls under the domain of sociology, although in recent years much of this work has been discussed in the human-computer interaction literature. The next several sections give an overview of some relevant aspects of this body of work.

1.5 Important Design Issues

There are a variety of concerns that will be important to address when building robots that are meant to interact with people in social ways. There are many important features of the robot that must be implemented well, such as the necessary agility to complete its tasks, robustness enough for the environment it should work in, or an appearance that is deemed acceptable to its desired audience. However, the aspects of social robots that we are concerned with here are mainly with regard to how they are perceived by people.

If a robot is designed to be depended on by a person for completing a particular task, it must be seen as trustworthy. This is not a feature that we know how to turn off or on, so it likely has something to do with how the robot interacts with a person. In interactions where a person is relying on a robot for information, they must believe that the information is credible. This is one of the features of robots that can be affected by various causes such as the presence or proximity of the robot [13, 43]. If this information is a result of the robot teaching a person, then the robot must be capable of engaging the person in that interaction. As with other aspects of social robots, there are many issues that affect engagement, as a number of studies have shown [13, 26, 42, 43].

For any type of robot that is will be used on a regular basis, it must be dependable. Being seen as dependable will help allow people to build up trust in the robot, to believe that it will be consistent in its operations, and that it will be available when it is needed. What a person thinks about the robot's motives is important as well. In many, although not all, of the types of interactions that were discussed previously, a robot that is perceived to have the person's best interests in mind is desirable. An altruistic robot will often be viewed as beneficial to someone interacting with it.

We know that having a person (or even a computer) near to us, rather than further away, is more likely to influence our decisions [7, 28, 29, 32]. Therefore it is likely that we want a robot to seem immediate and not remote when we are interacting with us. In tasks where the robot is conveying information to a person (such as giving directions, teaching a new skill, or collaborating towards a particular goal), it must be persuasive in its speech and believable to the person. As computers have been shown to often be seen as more persuasive than humans [5, 12, 13], this may be easy to accomplish.

2 Background

There are many studies that are relevant to this exploration of human-robot interaction. As the first chapter discusses, these mostly fall into two categories: studies conducted on human-human interaction (sociology) and those done on interactions between people and computers (human-computer interaction, or HCI). This latter category also includes a few studies that have been conducted in the area of human-robot interaction (HRI). This chapter presents an overview of those results that are most important to the work presented here and discusses how these impacted the planning of the two studies carried out as a part of this thesis.

2.1 Humans and Computers: HCI and HRI

There have been a number of studies conducted over the last several years on interactions between humans and robots. While there is not yet a comprehensive literature on all aspects of these interactions, there are several studies that are relevant to the studies conducted in the course of this thesis. Some of the more relevant findings are presented here as grounding for the design and expectations of the two experiments presented in the following chapters.

2.1.1 *The Media Equation* – inherent social responses to media

Sociable robotics is a rapidly growing area of interest to researchers. As such, it draws on related areas of work and brings them together into the context of building robots that will interact with people as partners in a variety of social situations. The work of Reeves and Nass [42] shows that even a minimal level of social cues present in an interface or technological artifact leads people to treat that system in a way psychologically similar to how they would treat another person. Many of the studies presented in their 1996 book were based on interactions between a human and a computer or a television. We believe that similar, but stronger, effects will be found when we look at interactions between humans and even simple robots. As robots become more lifelike, we expect that these effects will become even more pronounced.

As interactions between humans and technological artifacts, in this case robots, become more social, we also anticipate that these reaction will become more nuanced. This thesis also looks at some of the interactions between the type of task being performed, gender of the person in the interaction, and whether a robot is physically present, things that *The Media Equation* studies only begin to address.

2.1.2 Presence

A recent study completed at the NTT Communication Science Laboratories by Byron Reeves, Kiyoshi Kogure, and colleagues compared robots to animated characters on the screen [43]. In this experiment, participants interacted with either a small (eight to ten inch tall) robot that looked like a cartoon character rabbit or a video of the same robot shown on a television screen. Participants completed three different scenarios with the character: retail sales (the character tried to sell the participant a set of knives), nutrition and health (the rabbit asked participants questions about basic nutrition), and a reading survey (the robot inquired about the participant's reading preferences).

At the conclusion of the experiment, participants completed a questionnaire that contained measures of credibility and liking. Participants were also evaluated on their memory of the interaction by completing a recognition test. Finally, during the experiment, physiological data was collected to determine skin conductance level of the participant and skin conductance responses during each 30-second interval of the interaction.

The findings of this study were that when the robot is physically present instead of on the screen, participants liked it more, women judged it to be more credible, and men had a greater memory of the interaction. They also found that when the robot was shown on the screen, subjects showed a higher level of arousal as measured through skin conductance, which was contrary to their hypothesis. We measure similar responses of participants of the interaction in our experiments and will be interested to see whether our results agree with those shown in this study.

2.1.3 Mental models of robots

Other research has focused on the mental models that people develop when interacting with a robot, such as Kiesler and Goetz's 2002 study [25]. This study utilized scales to measure social and intellectual qualities of a person's mental model of a robot, but found that a better measure of anthropomorphism was still needed. They then adapted the Big-Five Inventory [23] to determine how people rated more or less "computer-looking" robots on this scale. The Big Five Inventory is designed to measure extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience of the participant completing the inventory. We chose scales, however, that were designed to measure another entity besides the person who is completing the inventory. They found that the presence of visible computer hardware decreased the rating of the robot on the Big-Five scales.

2.1.4 Robot personality

Goetz and Kiesler published an additional study at the same time as the previously mentioned one on the level of cooperation between a person and a robot that was supposed to encourage the person to perform a short exercise routine [21]. This study had participants interacting with a robot that exhibited one of two personality types: playful or serious. The playful robot joked with the participant and acted as though the exercise was fun, while the serious robot talked about the health benefits of the exercise task. Their results showcase an interesting contradiction that designers may well face when creating a robot intended to promote a serious task. Participants in their study liked the playful robot more, ranking it higher on personality scales, but they performed the desired task (completing zero to two minutes of exercise) for a longer period with the serious robot. While this suggests an interesting dilemma, longer studies will have to be carried out before we know if these effects will hold for longer than a several-minute interaction. A recent longer-term study has been carried out with users interacting with an animated relational agent [6], which showed that people felt better about their relationship with the agent when it tried to engage in relational dialogue with the person. Although it is likely that these effects will extend to robots (and potentially be more pronounced), this remains to be seen. Our desire is to build robots for a long-term interaction, so these are the kinds of studies that will need to be completed.

2.1.5 Appearance of robot

In a summary of the lessons learned from building robots that interact with people in social situations [8], Breazeal discusses creating robots in such a way as to take advantage of human social expectations. The idea behind this is a logical one that extends from research and observations of human interaction. It is still very difficult, if not impossible, to build a robot that can behave in a manner that is socially similar to an adult human. Humans, however, have a very different expectation of how an infant will respond to them. Thus, Breazeal states that we should build our robots with an appearance that will elicit the proper level of interaction from its interlocutor. In her previous work with the robot Kismet [11], she was careful to design the appearance of the robot and the set of interactions that it was capable of to reflect an infant-like quality. She was successful in encouraging people to use actions and vocalizations that are commonly associated with interacting with a small child or infant. This gave the robot's designers a simplified set of stimuli that had to be recognized by the system in order to provide a compelling and robust interactive experience between a human and the robot. This bootstrapping process is a fundamental part of current socially intelligent robots that allows them to take advantage of social developmental processes as we understand them to be taking place in a human infant.

2.1.6 Gender effects

The NTT rabbit study [43] demonstrates several effects that were before unseen in robots, although expected by the researchers. This study had users interacting with a robot that looked like a rabbit that was either physically present in front of them or displayed on a television screen. Participants were presented with three different scenarios (retail sales, nutrition and diet, reading survey) and answered a questionnaire on their experience after each interaction. One interesting effect shown in this study is that women liked the robot better than men, found it more credible, more like an equal, but remembered it less well. They caution that the robot (a stuffed rabbit) may have been inadvertently designed to appeal more to women, thus causing the more positive feedback about the robot. They note that the design of the robot is likely to be highly important in the resulting perceptions of those interacting with it and that these issues should be further studied.

2.1.7 Credibility

The Reeves, et al. study previously discussed [43] found that there was a significant effect on credibility caused by the presence of the robot, but only for women. When asked to rate how credible they found the robot and the on-screen character, women showed a clear bias towards the on-screen character; believing that is was more credible across all tasks. Responses given by men, however, showed almost no variation in credibility attributed between the physically present and the remote, televised robot.

2.1.8 Liking and Preference

Other questions in the Reeves, et al. rabbit study [43] asked participants about the general "likeability" of the robotic characters with which they interacted. Again, there were significant gender differences, but both men and women liked the physically present robot more than the one shown over the television. Responses from women showed a clearly larger difference (difference of approximately 4.5 for women versus approximately 1 for men on a 12-point scale) in the higher rating for the physically present robot. All of the results from this study lead us to expect significant gender differences in several of the scales being measured in this work.

2.1.9 Perceived information quality

A person's perception of the information that they receive from a computer is affected by their social identification with the computer, according to a study by Nass, Fogg, and Moon [37]. Their study had participants completing the Desert Survival Task with a computer, ranking the twelve items after reading about each of the twelve items on the computer and then exchanged information with the computer before making their final ranking. (The Desert Survival Task is explained in greater detail in the description of the second experiment in this thesis.) The independent variable in this experiment was what participants were told about their interaction with the computer. Half of the participants were told that they were working as a team with the computer, while the other half were told that they were working as an individual.

One of the indices measured at the conclusion of the experiment was the participants' perception of the quality of the information that they received from the computer. This index was made up of three items: the relevance of the computer's information, the helpfulness of that information, and the insightfulness of the information. (Cronbach's alpha is reported for this index as 0.92.) They reported one significant finding of this study being that participants in the team situation perceived the information quality from the computer to be higher than those in the individual situation ($t(26) = 2.96$, $p < 0.01$). If the social identification of the person with a computer can affect the perceived information quality, there could be other aspects of an interaction that may play a role in this as well.

2.2 Human Behavior: Social Psychology and Sociology

2.2.1 Presence

The terms social presence, telepresence, and simply presence have all been used in recent years to denote the idea of how closely a mediated experience is to an actual, "live" experience. Lombard and Ditton describe the range of characteristics that are meant by these terms in their 2000 work on measuring social presence [31]. In this paper, they also describe a scale they developed that can be used to measure presence.

This recent study follows from their extensive 1997 review of presence studies [30] in which they identify and explain the six conceptualizations of presence that have been used in media studies. Their scale combines several aspects of presence into one scale in an attempt to create a standard instrument for measuring presence. The six dimensions of presence that they incorporate are social richness, realism, transportation, immersion, a social actor within a medium, and a medium as a social actor. We chose to use this measure in our studies over the Big-Five or a similar scale because we view a robot as a new medium through which a message can be conveyed. This scale allows a participant in an interaction to rate the qualities of the other entity whereas the Big-Five is designed as a self-assesment scale of personality traits.

Lombard and Ditton's work to devise reliable and valid measures of the response of people to various mediated interactions can be used to study interactions between people and various types of robots. Although we are interested in how a robot compares to a human on a scale of social presence and believe that previous research suggests that in many ways people will treat a robot with the same social conventions that they would a human, our goal is not to construct a robot that is anthropomorphically attributed all aspects of a human's mental life.

Lombard and Ditton's six measures of are of interest in the study of human-robot interaction in the following ways: Presence as *social richness* comes from the development of social presence theory and media richness theory, which "were initially developed to better match communication media and organizational tasks to maximize efficiency and satisfaction." *Realism* is used to measure responses to changes in distinct aspects of the medium in question, such as the change in size of a television on which a particular video is watched. We are interested in measuring how people respond to changes in social cues presented by a robot during interactions. The third aspect of presence defined in this scale is presence as *transportation*. Lombard and Ditton provide several conceptual forms of this idea. The one most of interest to us is the idea of shared space. Although they present it in terms of presence through mediums such as teleconferencing and virtual reality, there are similar issues that are important in interaction through a robotic character.

Psychological and perceptual immersion is an important factor in determining how immersed a person can become with the medium. Presence as *a social actor within a medium* is defined by Lombard and Ditton as when "users respond to social cues presented by persons they encounter within a medium even though it is illogical and even inappropriate to do so." The studies presented in [38, 39, 42] suggest that even a minimal set of social cues will induce a person to use social behavior in response to a medium. Finally presence as *medium as social actor* differs subtly from the last conceptualization of presence presented in the case of robots. We may distinguish the two by noting that the social actor within the medium is the "content" contained within the medium and the medium as social actor is the medium itself. In the case of a robot, there is a high degree of interaction between these two cases, although on the same physical robot (medium), we can vary the social cues present (content) to elicit different response from a user.

Although we do not know of anyone using this type of measure to gauge the performance of a robot, we believe that this is a valid measure when comparing the performance of a robot to that of an animated character because the purpose of the scale is to quantify what they refer to as "the perceptual illusion of nonmediation." In other words, their scale measures how much like a direct, non-mediated interaction a given interaction with another entity seems to be. Although this other entity is typically a person, in the first experiment when we use this scale, the other is a person, a robot, or an animated character. As will be seen in the first experiment, we use only four of the six measures. The two not used are presence as *realism* and presence as *medium as a social actor*. This is further discussed in the next chapter.

2.2.2 Sociability

There have been several studies in recent years that have explored dimensions of interaction as they relate to mediated communication. Lombard and Ditton initially applied the work that was discussed in the previous section to the study of people's responses to various media [30]. Another study by Morkes, Kernal, and Nass [33], compared the responses of participants interacting with "another person in another room" through a computer-mediated communication (CMC) setup or interacting directly with another computer. (In actuality, participants were always interacting with a computer through the same interface; they were simply given different stories for whom or what was on the other end of the communication.) What they discovered was that people were more sociable when they thought that they were interacting with another person than with the computer. They also note that participants in the HCI case laughed and smiled significantly *less* that those in the CMC situation.

2.2.3 Social judgments

Another recent study by Burgoon and colleagues [13] found that the physical proximity of another person affects the judgments that are made about that person. This study had people interacting with an experiment confederate to complete the desert survival task (discussed in detail in the Second Experiment chapter). Participants interacted with the confederate face-to-face; at adjacent computers, but communicated only through text in an online chat environment; in separate rooms, communicating through text online; in separate rooms, communicating via audio conferencing on a computer; or in separate rooms and communicating through audio and video conferencing on a computer. The study found that "actual or perceived distance can indeed weaken… the credibility they ascribe to task mates." Thus even in a situation where the other interactant is another person, we see that closer proximity between the two participants in the interaction leads to greater credibility being ascribed to another participant.

This study by Burgoon and colleagues also found that when the other participant seemed to be closer to the participant, this led to higher social judgments being made about them (such as similarity to the participant, degree of sociability of the other participant, and trust of the other participant) as well as leading to greater interactivity between the two participants.

2.3 What Does This Mean for Robots?

Most of the studies discussed in the previous two sections never mention robots. A number of them never even address questions of technology, but focus on human-human interaction. So how do we get from these experiments to our current work in robotics?

Findings of previous interaction research show results in these categories:

Information factors

A *physically present* robot is seen as **more credible** than a *remote* robot.[1]

A task partner who is *closer* is seen as **more credible** than one *further away*.[2]

When a person feels a *closer affinity* to a partner, they believe their partner provides **higher quality information.**[3]

Task attributes

Proximity has no effect on **task performance.**[2]

Men remember more information from a *physically present robot*; **women show no difference.**[1]

Likeability

Humor causes people to **like a robot more.**[4]

A *physically present* robot is **liked more** than a remote robot.[1]

Social judgments

Women see robots as more of a social equal than men do.[1]

People make **higher social judgments** about a partner who is *physically closer*.[2]

Interactivity

Greater levels of interaction occur when partners are *physically closer* to one another.[2]

References

See entries at end of thesis for complete reference.

1. Reeves, et al. Robots Versus On-Screen Agents: Effects on Social and Emotional Responses.
2. Burgoon, et al. Testing the Interactivity Principle.
3. Nass, Fogg, and Moon. How Powerful is Social Identity?
4. Goetz and Kiesler. Cooperation with a Robotic Assistant.

Figure 1. Related research findings

We start with the thesis of *The Media Equation*, the Reeves and Nass collection of studies that give us a good picture of humans' reactions to and interactions with other forms of technology. This is combined with the work that was presented from the fields of sociology and social psychology, which allows us to say something about what we expect to find in interactions between people and robots.

Some of the findings about humans interacting with each other that concern proximity will likely extend to HRI: robots that are perceived as closer will be seen as more credible, more persuasive, and will score higher on judgments of a social nature. (See Figure 1 for a summary of the research discussed in this chapter.) We may also be able to extend the studies performed across multiple modalities of communication to predict that a robot will be found more engaging than an animated character, that a robot will be more persuasive when it is physically present, and that a greater level of interaction between a person and a robot will lead to greater trust and higher judgments of the performance of the robot by a person.

2.4 Overview of Work Presented Here

The studies that were undertaken as a part of this thesis address some of these important issues that come up when creating a robot that will interact in this manner. In the first study presented, we are particularly interested in how the robot performs with respect to the human and the animated character. We believe that the robot's physical presence will lead to a higher score than the animated character on measures of presence, realism, and other measured qualities of the interaction. We also anticipated that the robot will score lower than a human on these same measures. In the sections on the design of the questionnaire for the experiment, we discuss in greater detail the measures that were analyzed.

One limitation of the first study is that there is a confounding factor in the demonstrated difference in reactions shown between the interactions with the robot and those with the animated character. The question is as to whether the differences result from the animated character being not a real thing – i.e., something created solely within the computer while the robot is clearly a physical entity in front of the person – or that the animated character is not as physically proximate as the robot is. Stated simply, this comes down to a question of real versus fictional and present versus remote: do participants perceive a difference in the animated character because it's simply not in the same space (but *is* real) or because it is a fictional entity that is only portrayed on the screen?

The second study, in part, attempts to unravel this puzzle and make clear the distinction by testing only one of the two possibilities. In this experiment, half of the participants interact with a robot that is physically present and the other half work with the same robot that is presented to them on a television screen. Participants in both cases are aware that there is a real robot that is interacting with them, but only half of them see it before them.

The other aspect of HRI that is explored in this study is what effect the type of task has on a participant's perceptions and responses to interaction. Again, participants were divided into two groups with half performing a task in which they must cooperate with the robot and the other half learning a set of facts in a short lesson from the robot and asked a series of questions at the conclusion of the lesson. The features that were measured in this study included those calculated in the first study, allowing comparisons to be made with the earlier results.

3 First Experiment

The first study carried out as a part of this research was designed to elicit some of the differences that may exist when people interact with characters presented through different modalities. In this experiment, participants completed a simple task with a robot, an animated character, and a person. This chapter lays out the design of the experiment, considerations and options that went into the design, and the setup and protocol used. Later chapters present an analysis of the results and a discussion in light of other research.

The measures that were of interest in this experiment were aligned with two main questions. The first was whether there is a difference in arousal among the three interaction modalities. We are interested in knowing if participant's responses to a character differ depending on how that character is presented. The second question has to do with social presence and whether different levels of presence are attributed to each of the three characters. The definition that we are working with for each of these terms is discussed in the following section of this chapter.

3.1 Relevant Measures

3.1.1 Arousal

As was discussed in an earlier chapter, two of the more compelling applications of sociable robots in the future are in education and health care. We can imagine that in the near future robots could serve as assistants in a variety of situations in these areas. In educational applications, one of the most important elements of such a system is its effectiveness in teaching and for instilling the learner with the ability to later recall what she was taught. Reeves and Nass state that "it's the arousing experiences... that are best remembered. A growing literature shows that arousal, whether caused by positive or negative experience, may be equally (and some argue even more) important" [42]. This suggests that if we find a difference in the level of emotional arousal between two media experiences or interactions, then the one that brings about a higher level of arousal in the user may be better suited for an educational application. There is a limit to this theory that more arousal is better, and we will also have to work to understand when this limit is reached. With the small differences expected here, this should not be an issue, however.

The work of Dawson and others shows that certain human physiological responses are reliable and valid indicators of a person's level of emotional arousal [19]. We chose to use skin conductivity, an electrodermal response, as the physiological measure in our study because besides being a good indicator of motivational and attentional arousal, it is also unobtrusive to the participant and relatively simple to measure. This data will be analyzed for a mean skin conductivity level and the peaks that are expected to occur shortly after participants start a new activity with one of the characters.

3.1.2 Social Presence of Characters

As discussed in the Social Psychology section of the Background chapter, the scale developed by Lombard and Ditton for measuring four of the six conceptualizations of social presence [31] is the one used and reported in this study. Scores for all three characters used in the experiment are computed as one basis for understanding differences among people's perceptions of interactions within the three modalities.

3.2 Experimental Methods

3.2.1 Participants

The 32 participants ranged in age from 18 to 47 years of age ($\underline{M}$ = 27, σ = 9). Half of the participants were male and half were female. Eighty-one percent of the participants were white (n = 26), 9 percent were Asian (n = 3), and 3 percent indicated each of African American, Hispanic, and Other (n = 1 for each). Participants consisted of a mixture of graduate students, undergraduate students, and professionals from the local community. Because neither the robot nor the animation had been shown outside of our lab before the experiment, none of the participants had seen either before.

Participants were also asked to self-report scores on their knowledge of artificial intelligence and robotics on a seven-point rating scale (1 = none; 7 = a lot). The mean self-reported knowledge for robotics was 2.8 (σ = 1.6) and the mean self-reported knowledge of artificial intelligence was 3.0 (σ = 1.5). A higher number on both scales indicated a greater knowledge of the field. Although these results do not give us a definite evaluation of the knowledge of the participants, it does inform us that these participants were likely not experts in robotics or artificial intelligence; thus it is likely that they did not hold biased opinions about the research topic.

3.2.2 Experimental Design

Physical setup

The participant was seated across a table from three characters, a robot, an animated version of the robot, and a person (as shown in Figure 2). The participant was separated from each of the characters by a black screen with a rectangular cutout approximately 3 inches by 7 inches. This was done so that the participant could only see the eyes of each character, minimizing any effects that may be caused by differences in the technology (or lack thereof) surrounding the character. Thus the support structure and motors of the robot were hidden, the rest of the flat-screen monitor was concealed, and the participant could not see the remainder of the face of the human.

Figure 2. Three character modalities: robot, animated, and human

On the table between the participant and each character were placed a red, a green, and a blue wooden block, each approximately 2 inches square. (See Figure 3 for the layout.) The distance between the characters was roughly 18 inches. This setup allowed the participant to move after each of the interactions so that they were seated directly across for the character with which they were currently interacting. The other two characters were hidden from the participant during each interaction. The participant was also asked to adjust the height of the chair that they were seated in so that they were approximately at eye level with the character.

There were three video cameras set up to record the interactions for later analysis. The first was set up in front of the participant in order to record their facial expressions and movements. Another camera was placed behind the participant to record the actions of the character and the participant's movements with respect to the character. A third camera was set up at the right end of the table facing back across the table, placed vertically about halfway between the top of the table and the center of the eyes. This camera was in place to record the position of the blocks during the interaction and allowed us to determine when, where, and how each participant moved the blocks in response to the characters' requests.

Figure 3. Layout of blocks and character from subjects' point of view

Task

The task chosen for this experiment had the participant responding to spoken requests from the characters which asked the participant to manipulate the colored wooden blocks. As mentioned previously, one measure that the experiment design included is skin conductivity. A confounding factor of this measure is level of cognitive activity, so we deliberately chose a task that had a low level of cognitive engagement. For this study, the need for simple interaction between the participant and the character was more important than the nature of the particular task that was selected.

Commands spoken while looking at a particular block:

- Move this block towards me.
- Move that block off the table.
- Hold that block up so I can see it.

Commands spoken while looking at a point on the table:

- Move the blue block there.
- Put the yellow block here.

Commands spoken while looking at the subject:

- Move the red block towards me.
- Put the blue block where I can't see it.
- Please move the yellow block to my left.

Figure 4. Requests made by each character

There are nine requests that were made by the character to the participant during each of the interactions as shown in Figure 4. The requests were presented in a female voice and in a different order by each of the characters. (The choice of voice will be discussed shortly.) All of the requests require the participant to pick up and move one of the blocks and then replace it to its original position after a short pause. Once a request by a character has been made, there is a fixed, pre-determined amount of time before the follow-up appeal is made. The time between finishing one request and starting the next is not fixed and subsequent requests commence as soon as the participant has replaced the block from the previous action.

An example interaction is as follows:

- *Character*: "Move that block off the table." The character looks down at the red block on the table for approximately three seconds and then looks back up at the participant.

- The *participant* reaches for the red block and puts it in their lap or holds it up in the air.
- *Character*: "Thank you. You can put it back now." The character is looking at the participant while speaking.
- The *participant* places the block back on the table in the colored square where it belongs.
- After a brief pause, the character makes the next request to the participant.

After the nine short interactions have been completed, the next character is uncovered, the first one is covered, and the colored blocks are moved so that they are in front of the next character with which the participant will interact. The participant then repeats the series of requests, although in a different order, with the second character. Once he or she has finished interacting with the second character, the process is repeated with the final character.

Character depictions

The characters were designed so that the characters in different media would appear as similar to each other as possible to the participants participating in the experiment. The robotic eyes consist of a pair of eyes with two degrees of freedom each: left-right and up-down. This gives them a similar range and ability of motion as human eyes. Each eye also has upper and lower lids that can open and close.

The animated character was based on the robotic character and was created to look as similar as possible. The colors are matched, the movement is controlled in the same way as the movement of the robotic eyes, and their manifestation on the screen is such that they appear to be the same size as those of the robotic eyes.

Both the animated character and robotic character were controlled by the same computer program. This allowed them to act similarly to one another. The locations where they appeared to look, the timing of their blinking, and other small movements that they made were played out the same way on both characters. The human also appeared to look in the same directions as the other two characters and tried to do so with similar timing. The actions of the robotic and animated characters were based on the timings of the human in an attempt to make the movements realistic.

Figure 5. Human character

The human character was acted by the experimenter. As in the case of the other two characters, only the eyes of the human were visible. As can be seen in Figure 5, the rest of the face and body were hidden by the black screen and the table. In order to make the interaction as similar to the other two scripted interactions as possible, the experimenter chose locations around the room that were approximately the same as where the robot and the animated character appeared to be looking. At the appropriate times during the interaction, the experimenter looked in these locations. During the times when the other two characters were looking at the participant, the experimenter looked directly at the participant even if they were not directly in front of the experimenter.

We chose to create a character based only on the eyes for several reasons. With a simple character, it is easier to create the same character in different media that are similar to each other. With simply eyes, it is also more feasible to do a comparison with a human (with the rest of the face and body of the human hidden). With only eyes, it also becomes easier to attribute the findings of the study to the variable that we were changing: the modality with which the character was presented. A simpler character reduces potential confounding effects from participants' perceptions of the qualities of the character based on its appearance.

The differences between the interaction with a human in this experiment and a normal, everyday interpersonal interaction are obvious. However, it is difficult to make experimental comparisons between a robot of deliberately limited interaction capabilities and a human. For the measures that we were interested in comparing in this experiment, we decided that this setup would be the most useful. While we do make comparisons between reactions to the robot and to the human, we believe that more work should be done before more definite conclusions can be drawn from this information.

Voice of characters

The voice used to make the requests was a prerecorded human female voice. The same voice was used by all three characters to try to remove any effects of the particular voice from the results of the experiment. There were nine requests recorded as well as several responses that were played after the participant had completed responding to each request. (The requests are shown in Figure 4.)

Choosing the voice of the characters in this experiment was difficult and took a bit of consideration. The options for each character's voice span two choices, giving four options for each. The voice could be male or female and could be recorded human speech or computer-generated speech. Thus there were many choices for the overall combination of voice and character. Here we present several of the most strongly considered options and discuss why we made the decision that we did.

One option would be to have computer-generated speech for the robotic and animated characters and human speech for the human character. The argument for this combination of voices is that while one of the characters is clearly human and the other two are not, the voices should match those perceptions. The problem that this presents is that the effects from the different voices may overpower the effects of the different media that we were trying to measure. For that reason, we decided to use one voice for all three characters. The choice of a human voice over a computer-generated voice was made for two reasons. One is that recorded human speech is easier to understand than computer-generated speech, which makes the character easier to understand regardless of the modality. The other reason is that it would appear very strange to have the human character speaking to participants in a synthetic computer-generated voice. The cognitive dissonance created in this case would be likely to overwhelm other effects that we were trying to measure.

Another option would be to have either a male voice for all characters or some combination of male and female voices for the three characters. The decision to not use both male and female voices was made based on research that has shown that people perceive computers differently based on the gender of their voice, most notably work reported by Reeves and Nass on differing responses by people to computers which appeared to have voices of differing gender [42]. Because we were not studying the effects of gender, we chose to use a single gender for the voice across all characters to avoid effects of gender. In the end, the choice of a female voice over a male voice for the three characters was arbitrary. Although there were a multitude of options for the voices of the three characters and none of them without difficulties, we made the final choice of voice in order to minimize confounding effects on the experiment.

Within-subjects design

One of the measures that we were initially interested in was the level of arousal of participants across the different modalities as measured by their skin conductivity. Because it is known to be difficult to correlate skin response readings across multiple participants, we chose to implement this experiment as a within-subjects design. Although there is the potential to create demand effects on the participant which may affect the results, we believe that because of the short duration of the experiment and the simplicity of the task that these demands will not occur. The protocol experienced by each participant based in this design is discussed in detail in the following section.

3.2.3 Protocol

Each participant completed short interactions with each of the three characters: the robot, the animated figure, and the human. The duration of each of these interactions was approximately three minutes, allowing participants to complete the entire interaction portion of the experiment in less than ten minutes.

We took a Wizard of Oz approach [18] to the design of this experiment, allowing the experimenter to exert the necessary level of control over the order and timing of interactions during the course of the experiment. Using a prerecorded voice and preset timings for each exchange between a character and participant allowed us to insure that each participant would have the same experience.

The software to run the experiment was designed so that the experimenter could operate all of the necessary actions from one screen with either the mouse or the keyboard. This interface allowed the experimenter to remain out of sight while the participant was interacting with each character. During all three of the interactions, the experimenter was not visible. Between each of the interactions, the experimenter assisted the participant in moving before the following interaction and adjusted the placement of the experimental materials.

When a participant entered the room, he or she was seated in front of the three characters, all of which were visible. Participants were pre-assigned an order in which they interacted with the three characters. (All six possible orderings were used, with approximately one-sixth of the participants interacting with each ordering of characters.) After being seated, the participant was read a short introduction to the experiment. This introduction informed them that they would be interacting with the three characters and explained the task to them in the following way:

> You are being asked to interact with three characters. You can only see their eyes, but they are a robot, an animation, and a person. Each of these interactions will last only a few minutes and during that time, you will hear the character asking you to perform some simple tasks with the colored blocks that you see on the table.

> Please adjust your chair so that you are about at eye level with the characters.
>
> While you are performing these tasks, we would like to measure something called galvanic skin response using these two straps connected to two of your fingers. This measures how conductive your skin is, which changes depending on how interested and excited you are about what you are currently doing. It is not something that can easily be consciously changed. In a moment, I will put the sensor on and you can see how it feels. Please try to hold this hand still in your lap while you are doing the experiment.

The two characters not in use first were then covered and the participant was seated immediately in front of the character with which he or she was first interacting. The three colored blocks were placed in their home positions between the participant and the active character. These positions were denoted by a slightly lowered area in the tabletop in which the block fit. Each of the three square indentations were colored to match the color of one of the blocks. (See Figure 3.)

The interaction then proceeded as described in the Task section above. At the conclusion of the interaction with the first character, that character is covered, the blocks are moved so that they are in front of the next character, the participant moves so that they are in front of the next character, and the character is uncovered. The participant then completes a second interaction using the same requests presented in a new order. These steps are then repeated for the final interaction with the third character.

At the conclusion of the three interactions, participants were asked to complete a questionnaire about their experiences with the three interaction modalities. This questionnaire took participants approximately eighteen to twenty minutes to complete on average. It is described in detail in the next section of this chapter.

3.2.4 Dependent measures

The questionnaire given to the participant at the conclusion of the interactions consisted of four distinct sections. These sections were designed or adapted to measure several characteristics of the interactions between the participants and the three characters. Questions in the first two sections were taken from an earlier questionnaire [31] which was designed to measure the five of the six dimensions of social presence discussed earlier. These first two sections also measured level of engagement of the participants. The third section was comprised of a set of adjectives that we have added to describe the characters and the interactions. The questionnaire concluded with a final section consisting of several open-ended questions to elicit further information from the participants. These questions asked the participant to report what they thought was missing from each character and which one they enjoyed interacting with most. Participants were asked a total of 35 questions about the interaction and rated the characters on 30 adjectives. The questionnaire concluded with several questions to gather biographical and background data on the participants. The entire questionnaire is reproduced in Appendix A.

3.3 Results

The data collected from this experiment was analyzed with respect to the several questions that were laid out at the beginning of this chapter. Data from two participants was removed before analysis began because all answers were identical after the first page of the questionnaire. This left data from 32 participants. In some cases, participants chose not to answer particular cases. These are noted in the following sections where n is less than 32.

3.3.1 Qualities of the Interactions

Participants were asked to read and evaluate a series of statements and questions about engagement with the character on a seven-point scale. Six of these statements are presented in Figure 6 along with the average score for each character by the participants. In all cases, n = 32.

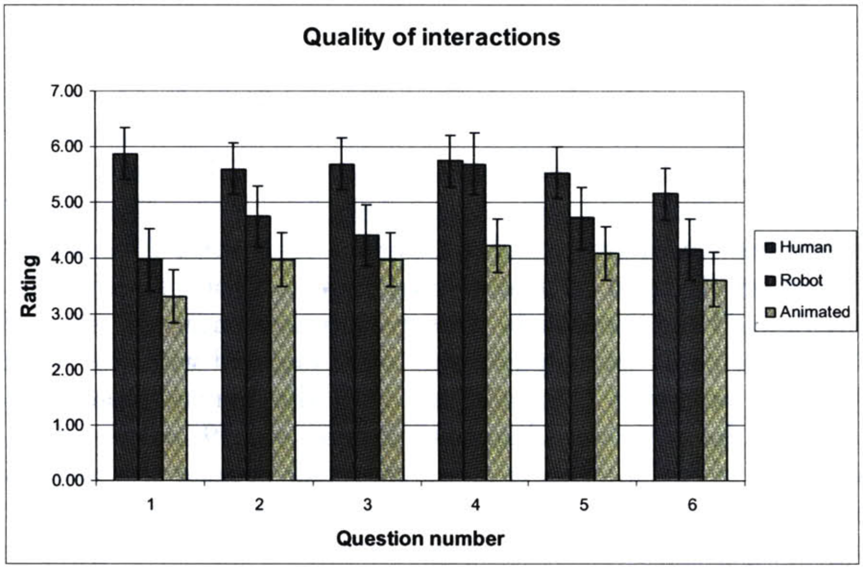

Figure 6. Quality of interaction items

Results from quality of interaction questions shown in Figure 6 where subjects are responding to the following questions: (All responses were on a 7-point scale with a higher number corresponding to more affirmative response. For all questions, $p < 0.001$.):

1. How often did you feel that the character was really alive and interacting with you? ($F(2,93) = 37.69$)
2. How completely were your senses engaged? ($F(2,93) = 22.19$)
3. To what extent did you experience a sensation of reality? ($F(2,93) = 12.09$)
4. How well were you able to view the character from different angles? ($F(2,93) = 16.74$)
5. How engaging was the interaction? ($F(2,93) = 8.98$)
6. The experience caused real feelings and emotions for me. ($F(2,93) = 16.21$)

As was anticipated when measuring aspects of lifelike interactivity, the human usually comes out most highly ranked. As can be seen in our data, this is indeed the case. When asking participants to evaluate qualities of their interactions with the characters, one with the greatest difference was "How well were you able to view the character from different angles?" The results gave ratings of 5.74 for the human character, 5.67 for the robotic character, and 4.22 for the animated character. This result can be expected since we are comparing two three-dimensional characters (the human and the robot) to a three-dimensional animation displayed on a two-dimensional computer screen. The questions "How completely were your senses engaged?" showed that participants were significantly more engaged with the robot than the animated character and with the human more than either of the other two characters (human = 5.59, robot = 4.75, animated = 3.97). This is as we expected: because this scale is measuring how much like a normal human-human social interaction these interactions were, a human should rate higher on the scale than the characters in other modalities. We hypothesized that the robot would be the next highest ranked, followed by the animated character with the lowest ranking.

Several questions related to how "alive" and "real" the characters appeared to the participants during the interactions. Questions 1, 3, and 6 in Figure 6 show a similar pattern of responses to these types of questions. In all three cases, the human had the highest rating, followed by the robotic character, then by the animated character. The differences shown in the table above indicate the differences between the human and the robot are approximately two to three times the difference between the robot and the animated character. The only question in this section where the human scored the lowest was when the participants were asked, "How much attention did you pay to the display devices/equipment rather than to the interaction?" We believe that this is because, by habit, people are used to interacting with other people and not robots or characters on a computer display, so they only paid attention to the novel characters, the physical robot and the animated character on the computer display.

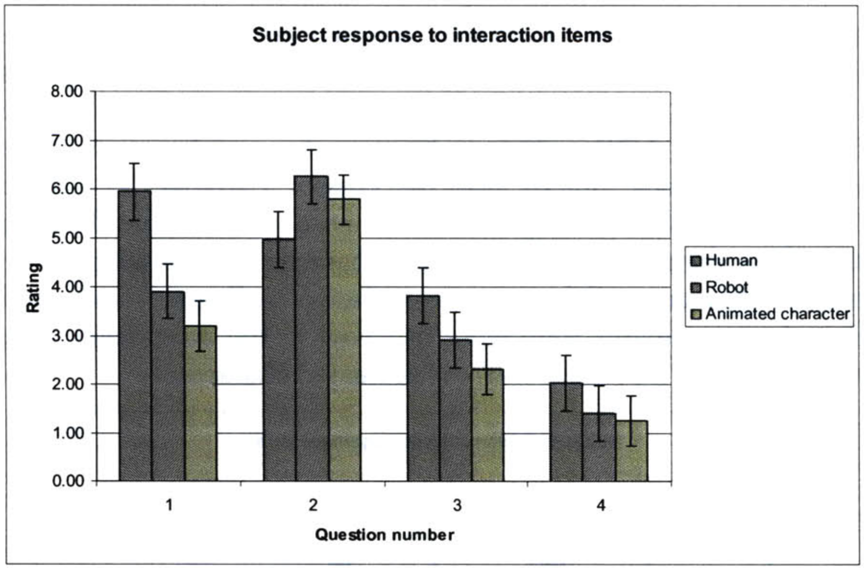

Figure 7. Interaction item responses

3.3.2 Participant Responses to Interactions

Participants were asked to rate how they responded to the three characters during the interactions. Some of the responses from this section are shown in Figure 7. The responses are all on 7-point scale with higher number corresponding to more affirmative response and correspond to the following survey questions: 1. How often did you have the sensation that the character could also see/hear you? ($F(2,93) = 31.41$, $p < 0.001$) 2. How often did you want to or did you make eye contact with the character ($F(2,93) = 7.55$, $p < 0.001$) 3. How much control over the interaction with the character did you feel that you had? ($F(2,93) = 12.84$), $p < 0.001$) 4. How often did you make a sound out loud in response to someone you saw or heard in the interaction? ($F(2,93) = 8.39$), $p < 0.001$)

One question about their perception of the characters' abilities, "How often did you have the sensation that the character could also see/hear you?," showed that the human rated much higher than the other two characters. The difference shown between the robotic and animated characters is still significant (two-tailed t test yields $t(31) = 2.26$, $p < 0.05$).

One set of responses that does not follow the usual ranking of human followed by robot followed by animated character are those to the question "How often did you want to or did you make eye contact with the character?" In this case, the ordering was robot ($\underline{M}$ = 6.25) followed by screen ($\underline{M}$ = 5.78) and by the human ($\underline{M}$ = 4.97). This is because, by social custom, it is unnatural to look directly into the eyes of a stranger at a close distance, so the participants were more comfortable looking directly at the robot or the animated character than looking at the human character. Other responses shown in Figure 7 indicated that the participants were more involved in the interaction with the robot than they were with the animated character.

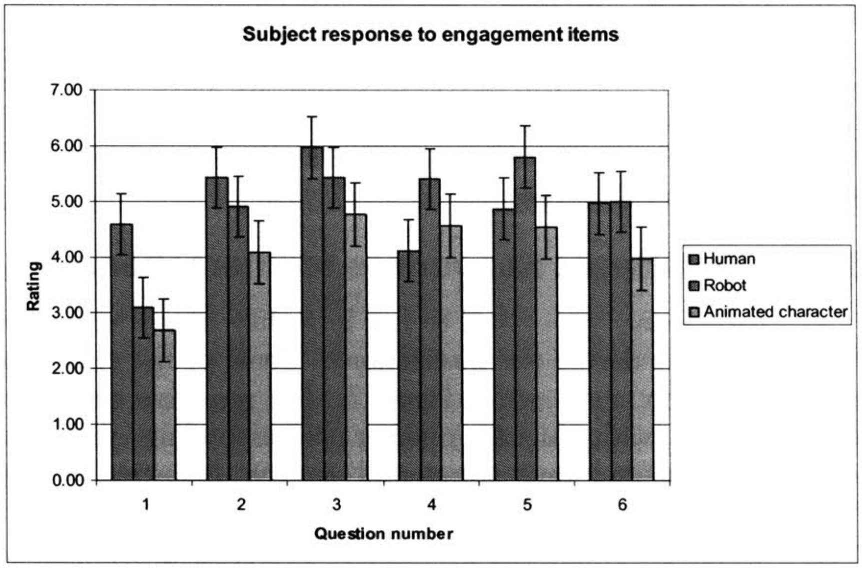

Figure 8. Engagement item responses

3.3.3 Engagement with Characters of Different Media

Several questions and statements were directed at eliciting the participants' feelings on how involved they were with the characters and how much desire they had for further interaction with each of the characters. Figure 8 presents results from this section of the questionnaire. The first three statements in that table showed a greater difference between the robot and the animated character than between the human and the robotic character. (These statements are: 1. He/she is a lot like me. ($F(2,93)$ = 14.14, p < 0.001) 2. If he/she were feeling bad, I'd try to cheer him/her up. ($F(2,93)$ = 7.61, p < 0.001) 3. He/she seemed to look at me often. ($F(2,93) = 7.16$, $p < 0.01$).) This indicates that the robot is seen to be more similar to the human than to the animated character. For the following three statements, the robot is scored higher than the human (significantly for the four and five, not for six). (These statements are: 4. I'd like to see/hear him/her again. ($F(2,93) = 9.92$, $p < 0.001$) 5. If there were a story about him/her in a newspaper or magazine, I would read it. ($F(2,93) = 6.80$, $p < 0.01$) 6. I would like to talk with him/her. ($F(2,93) = 7.26$, $p < 0.001$).) We believe that this is because of the unnaturalness of the close interaction between the participant and a person with whom they are not familiar, again as a result of social custom dictating that there is naturally a greater distance kept between strangers. This shows us that although people often treat robots and animated characters in social ways as they would another person, some of the social constraints that are present in interpersonal interaction do not come into play during interactions with a non-human.

Six of the statements regarding engagement in the interaction are combined into Lombard and Ditton's engagement scale. The responses for this scale are shown in Figure 8. When we compare robots to humans here, we can see that humans are seen as more engaging, as we expected. It is worth noting that this happens even though there is a mismatch between the female voice and the male face. The difference here, however, is not strongly significant (one-tailed t test gives $p < 0.06$, $t(61) = -1.58$). When we look at the difference between the robot and the animated character, however, we see that the robot is more engaging than the animation, and that this result is near significant (two-tailed t test yields $p < 0.08$, $t(62) = 1.76$).

3.3.4 Description of Characters

The final section of the questionnaire asked participants to rate how well a list of adjectives described each of the three characters on a scale from one to seven. The seven adjectives that showed the greatest differences are shown in Figure 9. (p-values were calculated using a within-subjects, repeated-measures ANOVA.) This set of descriptive adjectives highlights areas in which participants perceive a difference among the characters. We can see that the robot is seen as more convincing than the animated character ($\underline{M}$ = 4.25 for robot and $\underline{M}$ = 3.56 for animated character) and that both score below the human ($\underline{M}$ = 5.16). Although the actions of both the robot and the animated character were programmed to be identical, participants reported that the robot was more varied in its actions than the animated character ($\underline{M}$ = 3.45 for the robot versus $\underline{M}$ = 2.90 for the animated character). A similar response was seen for which was more entertaining (robot $\underline{M}$ = 5.41 and animated character $\underline{M}$ = 4.72) and compelling ($\underline{M}$ = 4.56 for the robot and $\underline{M}$ = 3.84 for the animated character).

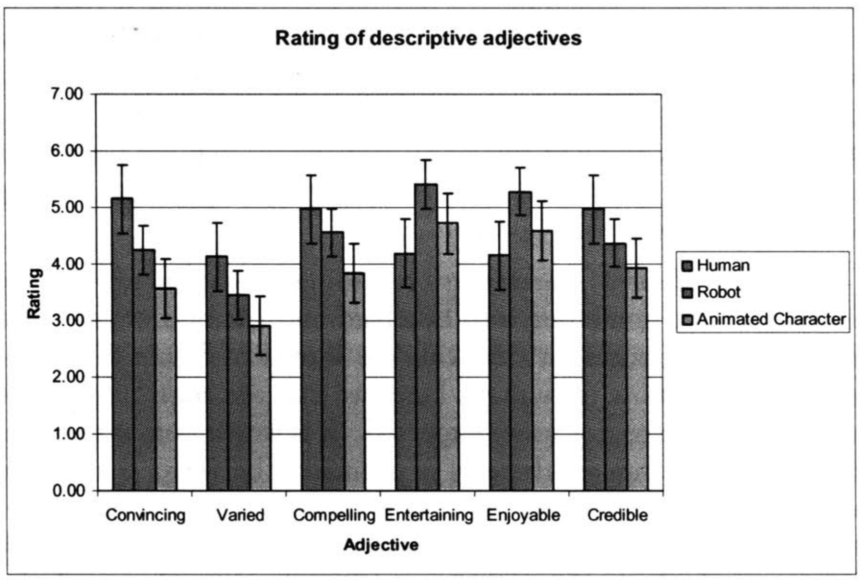

Figure 9. Responses to adjectives describing characters

Combining several of the adjectives into a scale asking how much participants liked the interactions across the different modalities, we see that the interaction with the robot is preferred to the interaction with the animated character, although this result is not highly significant ($F(2,93) = 2.72$, $p < 0.7$). This scale includes the adjectives annoying (which was reverse coded), compelling, convincing, enjoyable, entertaining, and likable.

We found an unexpected result in the cases of "entertaining" and "enjoyable," where both the robot and the animated character scored higher than the human. As discussed above, we believe that this is because of the participants being uncomfortable with interacting with a stranger at close range and as a result of participants attributing a richer inner psychology to the human than to the robot or the animated character.

3.3.5 Social Presence

Lombard and Ditton's six measures of social presence were discussed in detail in the Background chapter. Here I report participant responses to four of those scales that were measured using the questionnaire in this study along with a brief recapitulation of their meaning and importance to the study of human-robot interaction.

Presence as transportation

Presence as transportation refers to three closely related aspects of perceived collocation of the person and the character, although Lombard and Ditton delineate each of them. These are the concepts of "you are there," "it is here" (referring to the character), and "we are together." All of these concepts have to do with sharing the same space: the first places the participant in the character's environment; the second is the feeling of the character being in the participant's environment; and the third refers to both participant and character seeming to be in the same place, but not addressing where that space is. For purposes of this study, we did not distinguish among the three cases.

What we discovered is that in comparing the robot to either the human or the animated character, there is a significant difference in participant responses to these items. The participants clearly had the strongest sense of shared space with humans, as would be expected. A within subjects, repeated-measures ANOVA shows a significant difference among the characters ($F(2,93) = 26.86$, $p < 0.001$; Human $\underline{M} = 40.8$, $\sigma = 7.0$; Robot $\underline{M} = 36.1$, $\sigma = 7.6$). If we look just at the difference between the human and the robot, we find a significant difference as well

There is a definite difference in the perceptions between the robot and the animated character as well. (Animated character $\underline{M} = 30.8$, $\sigma = 9.0$; $t(60) = 3.56$; $p < 0.001$) This tells us that while a robot is not seen as sharing the same space quite as the human is, it is definitely viewed as more similar to the human in this respect than the animated character is.

Presence as social richness

The concept of presence as social richness refers to a medium being seen as "sociable, warm, sensitive, personal or intimate when it is used to interact with other people." [31] As in the previous section, the ordering of responses was human, then robot, followed by animated character. The ANOVA shows a difference among the cases: $F(2,93) = 4.04$, $p < 0.05$. In this case, however, there was not a significant difference between the scores of the human and the robot ($p < 0.15$ with a two-tailed t test, $t(62) = -1.48$). The difference between the robot and the animated character was also not significant, but the gap between the human and the animated character was large enough to show significance. (Again, with two-tailed t tests, we see: between robot and animated character $t(62) = 1.30$ with $p < 0.20$ and between human and animated character $t(62) = -2.94$, $p < 0.01$. Averages and deviations for all three modalities are: human $\underline{M} = 30.3$, $\sigma = 10.7$; robot $\underline{M} = 26.7$, $\sigma = 9.5$; animated character $\underline{M} = 23.8$, $\sigma = 10.4$.)

Presence as immersion

Immersion as one of the conceptualizations of presence means that the participant becomes involved and engaged in the interactive experience. In our measurements of the four conceptualizations of presence presented here, we found that the most immersive experience was with the human ($\underline{M} = 34.0$, $\sigma = 8.6$), the second most immersive modality was the robot ($\underline{M} = 30.8$, $\sigma = 10.0$), and finally the animated character ($\underline{M} = 26.9$, $\sigma = 10.1$). Within subjects, repeated-measures ANOVA yields $F(2,93) = 9.12$, $p < 0.001$.

The difference between human and robot was not significant (two tailed t test yields $t(61) = -1.75$, $p < 0.10$), although the difference between the robot and the animated character was ($t(62) = 2.72$, $p < 0.05$). Although not every case of comparisons of these results were statistically significant, it is interesting in that it follows the general trend that we have seen throughout most of the results of this experiment.

Presence as social actor within medium

Finally, presence as social actor within a medium means that people will respond to social cues presented by the character even though the character is remote and cannot respond in kind. In this experiment, based on self-reported answers to questionnaire items, we found that there is virtually no difference between the human and the robot (human $\underline{M} = 52.4$, $\sigma = 15.6$; robot $\underline{M} = 53.4$, $\sigma = 13.3$). The animated character, however, was rated somewhat lower on this scale ($\underline{M} = 46.4$, $\sigma = 16.0$). The ANOVA shows that there is a significant difference here. ($F(2,93) = 4.28$, $p < 0.05$) The difference between either the human and the animated character or the robot and the animated character were nearly significant ($t(62) = -2.4$, $p < 0.05$ for the former, while $t(62) = 2.9$, $p < 0.01$ for the latter using two-tailed t tests.)

This difference appears to be a result of participants' beliefs that the robot could see and respond to them during the experiment. Although the robot did not in fact have this ability, numerous participants commented at the conclusion of the experiment that the robot could follow their movements or would react to their actions, while they did not state this belief about the animated character.

Presence as realism and presence as medium as social actor

These final two measures of social presence that were discussed in the Background chapter were not used in this study. The first, presence as realism, refers to changes within the medium. As this is the independent variable in this experiment, it is not being measured. Presence as medium as social actor explores how the medium itself is perceived as a social entity. As noted previously, this measure is quite similar to the conceptualization of presence as a social actor within a medium, which is what we have chosen as a more relevant measure for this study.

3.3.6 Electrodermal Response

The electrodermal response data that was collected was analyzed for skin conductance responses and average skin conductance levels during the experimental trials. This data was compared across modality differences, order of character differences (seeing robot before versus after animated character, for instance), and order of stimulus differences (comparing first response to each of the three characters across participants, for example). Unfortunately, none of the data was significant and is therefore not discussed further.

There are several possible reasons why this data is not usable for purposes of evaluating the outcome of this experiment. One is that the process of collecting it often introduces a great deal of noise into the data. The setup used in this experiment required participants to attach straps with electrodes around two of their fingers on their non-dominant hand. They were asked to keep this hand still in their lap while completing the interactions with their other hand. Even small movements of this hand introduced artifacts into the data that can overwhelm the actual signal that we are trying to gather.

Another potential problem is related to the characteristics of the human skin conductivity response. After a stimulating event occurs, it can take from several seconds to several minutes for the participant's skin conductance to return to its baseline level. This experiment was not effectively designed to measure skin conductance because of the close timing (fifteen to twenty seconds) between successive events. Thus it may be difficult to discern the effect of a particular stimulus (either request or modality) from the effects that preceded that point in the interactions.

Finally, the effect that we were looking for was one caused by the three characters. It is also possible that the effects of the task were stronger than those caused by the characters. There may be a difference in skin conductivity levels exhibited in interactions with the different characters. If this difference is small, it is not unlikely that the responses exhibited in response to the task may overwhelm the character differences, thus making it difficult to ascertain character differences from the data collected.

3.3.7 Video Tape Data

One measure that we hypothesized would be different across the three modalities is the ease with which participants could detect which block the character was looking at. This data was analyzed in two ways. The first was by looking at the time it took for participants to respond to the first request by each character that required them to determine which block to pick up by where the character looked. (As you recall from the earlier description of the task, some of the requests were worded similar to "Pick up that block and hold it where I can see it," requiring the subject to determine which block the character was referring to by where it looked. Other requests were explicit about which block was being requested and were worded like "Hold the blue block up to my left." In both cases, the robot would look at the block which it was requesting; the only difference was whether there was an explicit linguistic cue.) Because of the amount of variance in the response times, there was no significant difference found. (ANOVA yields $F(2,82) = 1.05$, $p < 0.36$. Data for all were: human: $\underline{M} = 2.52$, $\sigma = 0.18$, robot: $\underline{M} = 2.19$, $\sigma = 0.14$, and animated character: $\underline{M} = 2.35$, $\sigma = 0.19$.)

We also looked at average times over all of the interactions between participants and each character where the character indicated which block by looking at it. This test also proved inconclusive. In this case, the ANOVA gave $F(2,272) = 0.74$ with $p < 0.48$. (Data for all were: human: $\underline{M} = 2.41$, $\sigma = 0.07$, robot: $\underline{M} = 2.30$, $\sigma = 0.06$, and animated character: $\underline{M} = 2.31$, $\sigma = 0.08$.)

These results may come from the fact that the movement of the robot being extremely limited. Because we were trying to control all of the conditions to make them as similar to one another as possible, this restricted the natural freedom that the human and the robot might gain in an unrestricted interaction. Further work will have to be done to see if the robot would be easier to view than the animated character.

3.3.8 Robot as Compared to Other Modalities

In general, most of the data shows the human being preferred to the robot. It also shows that the robot is, in turn, preferred to the animated character. This is the case for engagement, likeability, and overall quality of the interaction. The data follows the trends that were expected in the hypotheses for this experiment.

Two results that were contrary to expectations were those for how entertaining the characters were and how enjoyable the experience of interacting with each character was. The robot compared favorably to the animated character, as hypothesized. Both of these characters were rated more highly than the human, which runs counter to our hypothesis. Potential explanations for this outcome are that the awkwardness of interacting with a person with whom participants were not acquainted was much greater than that experienced when interacting with the robot or the animated character. Another reason for these results are that participants experienced cognitive dissonance when experiencing the female voice apparently emanating from the male face.

The difference that is most striking in this data is the perceived difference between the robot and the animated character, especially on the scales of engagement and several of the conceptualizations of social presence. One thing that is unclear from looking at this data is why that is the case. We cannot be certain where the difference comes from. Is it the fact that one is physically present, while the other seems to be remote and shown only on a screen? Or is it that the robot is perceived as a real thing, while the animated character is a fictional entity simply portrayed on the computer display?

3.4 Summary

This experiment clearly showed that there are differences among the three modalities compared. People do not always treat an interaction with a human, a robot, or an animated character in the same way. We also can conclude that interactions with robots can be construed somewhere between interactions between people and interactions between people in computers. This puts us squarely in the middle of the sociology and HCI camps. We must, however, continue working to understand the cause of some of these differences to understand how they should impact our creation of social robots.

4 Second Experiment

One of the significant results from the first experiment was that there was a clear difference in participants' reactions to the robotic character and the animated character. Although this is an interesting finding and tells us that there is a difference between these two media, there is more than one reason why this disparity in reactions may occur. One cause of this could be that people see the robot as a real entity because it is physically in front of them, while they perceive the animated character as something that is not real, or fictional, because it is shown only on the screen. Another possible cause is simply the physical presence of the character. In the latter case, the robot would still be viewed similarly to the former, in that it is physically present. The difference is that the animated character could be perceived as though it is something that is real, simply not physically in front of the viewer.

A short way to sum up this problem is with the phrase used by Cliff Nass to refer to it: the "real versus fictional" question [36]. In other words, are the animated characters seen as real, but not present, or are they viewed as not real and simply a fictional character portrayed on the screen?

To address this question, the second study was designed to test a different possible answer to this question than the first experiment. In the second experiment, I control for the presence variable. This is accomplished by having half of the participants interact with the physically present robot and the other half interact with the same robot presented on a television screen instead of sitting directly in front of the person. Whereas in the first study, participants interacted with a robot and an animated character (the fictional case), they interacted with a robot or a televised robot in the second experiment (the real case).

Another question not answered in the first experiment is how the task that the participants participate in changes their responses to the robot. In the previous experiment, all participants completed the same task – moving blocks around at the request of the character. In the second experiment, participants completed one of two tasks. One was an interactive, cooperative task where the participant worked with the robot to solve a problem. The other task was less interactive and consisted of the robot reading a short lesson to the participant and then quizzing them on the lesson. Both are described in more detail later in this chapter.

The remainder of this chapter is laid out similarly to the previous chapter. I first introduce the specific studies and measurements from the literature that are important to this work. I then discuss several aspects of the design, setup, and protocol for this experiment. The chapter concludes with a short note on some of the resulting measurements. The following chapter then discusses the results from this and the previous experiment in greater detail.

4.1 Relevant Measures

4.1.1 Trust

One of the important issues in many of the interactions that are envisioned between robots and humans is the level of trust that the person has for the robot. In any situation where a person relies on a robot (for information or to complete a task, for example), the interaction will be better when the person can trust the robot to act as expected.

To measure the trust felt by the person towards the robot, the receptivity/trust subscale of the Relational Communication Scale developed by Burgoon and Hale [44] was adapted to apply to the robot. (The adaptation in this case was simply changing "he/she" in the original scale to "the robot.") This scale consists of six questions that inquire about several aspects of the participants' beliefs on the openness and sincerity of the robot.

The entire scale is reproduced in Appendix B. Cronbach's alpha for this scale was calculated to be 0.79. This calculation and all others reported below for Cronbach's alpha were done on the data that is presented in this chapter.

4.1.2 Perceived Information Quality

In interactions where the robot is a source of information for the person interacting with it, the person's beliefs about the quality of that information is important. Particularly in educational applications, the learner must be believe that the information coming from the robot is reliable and true.

In order to measure this element of the interaction, we used the perceived information quality questions from Nass, Fogg, and Moon's study on affiliation effects between people and computers [37]. These questions ask about the usefulness of the robot's suggestions in the desert survival task. These questions are shown in Appendix B. Cronbach's alpha for this scale was calculated to be 0.77.

4.1.3 Altruism

A person's beliefs about a robot's intentions may have an important bearing on how much they trust and can therefore work with the robot. For this reason, we chose to develop a scale to measure the altruism of the robot. While similar to the trust of the robot, this scale asks more about the perceived motivations of the robot than the trust scale does.

The scale consists of three questions (shown in Appendix B) that ask participants about the intentions of the robot. Cronbach's alpha for this scale was calculated to be 0.82.

4.1.4 Engagement

Knowing how engaged a person is in their interaction with the robot is an important indicator of the success of the robot in drawing the person in. According to Lombard and Ditton [30], when a participant is more engaged, or involved, the experience is "a direct and natural experience rather than just the processing of symbolic data and is therefore likely to be more compelling." They go on to state that in their experience, "although part of the involvement effect is likely due to the interactive, and therefore active rather than passive, nature of high-presence media, there seems to be more at work." Thus we are interested in measuring the level of engagement across the various cases of this experiment, both for direct comparison to one another and for comparison to the levels measured in the first experiment. Cronbach's alpha for this scale was calculated to be 0.71.

4.1.5 Reliability

The perceived reliability of a robot is potentially an essential element related to how much a person will be willing to depend on the robot in an interaction. If there is a low expectation of reliability, people will be unlikely to expect the robot to perform well on a task or to be available to perform the task when needed. We wanted to understand whether the perceived reliability of the robot was affected by the factors that varied in this experiment: presence and type of task. Cronbach's alpha for this scale was calculated to be 0.85.

4.1.6 Immediacy

Behaviors that bring people physically or psychologically closer together are known as immediacy behaviors. Although initially construed as a measurement of interpersonal behavior, this measure has a long history of being used in instructional communication. As one of the tasks in this study is instructional in nature, we believe that it is informative to measure the immediacy felt by the participants towards the robot. There are behavioral aspects of immediacy that can be measured (for example, positive head nods; direct, relaxed, and open-body positions, and close physical distance) as well as a questionnaire-based measure. In this study, we chose to use one of the behavioral aspects, physical distance, as well as the questionnaire developed by J. F. Andersen and reported in *Communication Research Measures* [44]. Cronbach's alpha for this scale was calculated to be 0.94.

4.1.7 Credibility

Another issue that is relevant in rating the performance of a robot after a task in which it conveys information to a person is how credible the robot seems to be. While the perceived information quality scale measured what participants thought about the particular information that was given to them by the robot, this scale is used to measure how believable the robot is as a source of information. Although there are no reported measures using this scale to measure human performance on tasks similar to those used in this experiment, we can use it to get a measure of the differences found among the experimental conditions tested here. Cronbach's alpha for this scale was calculated to be 0.85.

4.1.8 Persuasiveness

In the cooperative task used in this study, participants were given the opportunity to let the robot convince them that its answers were better than their own. Because there were no *a priori* correct answers, we measured the degree to which participants conformed to the robot's answers and used this behavioral appraisal to judge the persuasiveness of the robot. For applications of robotics in which the robot is intended to convince or inform a person on some topic, the ability of the robot to persuade is important. We hypothesize that when the robot is physically present, it will be viewed as more persuasive than in the telepresent case.

4.2 Experimental Methods

4.2.1 Participants

We recruited 82 participants for this experiment, giving approximately 20 participants in each of the four experimental conditions (robot present, cooperative task; robot present, teaching task; robot remote, cooperative task; and robot remote, teaching task). Participants ranged in age from 18 to 61 years of age ($\underline{M}$ = 27.7, σ = 9.7). Fifty-nine percent of the participants were male (n = 48) and 41 percent female (n = 34). Seventy percent of the participants were white (n = 57), 20 percent were Asian (n = 16), 2 percent African American (n = 2), 1 percent Hispanic (n = 1), and 7 percent did not report their race (n = 6).

Participants were recruited from the institute and the surrounding community. Approximately 63 percent of the participants (n = 52) reported their occupation as student, with the rest of the participants coming from the surrounding community. As in the last experiment, the robot used in this experiment was built expressly for this study, so none of the participants had seen it before.

Also as in the previous experiment, participants were asked to self-report scores on their knowledge of artificial intelligence and robotics on a seven-point rating scale (1 = none; 7 = a lot).

The mean self-reported knowledge for robotics was 3.87 (σ = 1.85) and the mean self-reported knowledge of artificial intelligence was 3.85 (σ = 1.79). A higher number on both scales indicated a greater knowledge of the field. When compared to the previous study, participants ranked themselves as more knowledgeable of both robotics and artificial intelligence. Although these results do not give us a definite evaluation of the knowledge of the participants, it does inform us that on average these participants were likely not experts in robotics or artificial intelligence.

4.2.2 Experimental Design

Physical setup

For this experiment, participants were seated at a table in an office. On the other side of the table, directly across from the participant, was either the robot or a twenty-inch television turned on its side. As discussed previously, half of the participants interacted with the robot physically present, while the other half viewed the robot on the television screen. The television was placed so that it was in the plane that the robot occupied when it was standing straight up, as seen in Figure 10. Thus for half of the participants, the robot could actually lean out across the table towards them, while for the other half it only appeared to do so.

When the robot was not physically present, it was set up in the room next to the room where the experiment was taking place. There was a video camera placed in front of the robot at the appropriate distance and orientation so that the robot on the screen appeared to be the same size and distance from the participant as the real robot would. One of the detrimental aspects of the robot that was built for this experiment is that the inexpensive servos used to move its joints were rather noisy. To insure that this noise would not be a factor that would make a difference in the present versus the remote interactions, the volume on the television was adjusted so that the noise was as audible to the participants who interacted with the robot through the television screen as it was to those who interacted with the robot directly in front of them.

For all participants, there was a flat screen, eighteen-inch computer monitor to their right side on the table during the interactions. During the desert survival task, this is where the task was displayed, while the monitor showed two maps of the area being discussed during the teaching task. There was a computer mouse on the table during the tasks that the participant could use to interact with the simple on-screen interface to complete the desert survival task or to commence the teaching task. A computer keyboard was not present during the interactions.

Desert survival task

The cooperative task chosen for use in this experiment is the desert survival task. This task was developed by Lafferty and Eady [27] and was used by Nass, Fogg, and Moon in their study of the affiliations between humans and computers [37]. The task is useful for this kind of interaction because it requires users to consider the validity of the answers of the other person, computer, or robot with which they are interacting. Because this task has been successfully extended from use in human-human interaction to human-computer interaction, we believe that it will also be a useful task in the present experiment.

This task presents participants with a list of twelve items that they might find useful if they were stranded in a desert and asks them to rank them in order of importance. Participants do this twice – once on their own and once in collaboration with the robot. This allows us to measure how much the participant conforms with the robot's response to the same task. The protocol for the interaction will be discussed later in this chapter.

Teaching task

The teaching task is very different than the desert survival task. Besides being a different nature of interaction, there is much less interactivity between the robot and the participant in this task. In this task, the participant is taught a lesson by the robot through the robot reading the text aloud. During this teaching phase of the experiment, the participant passively listens to the robot. After the lesson has been read, the participant is then asked a set of questions by the robot and is expected to answer them out loud.

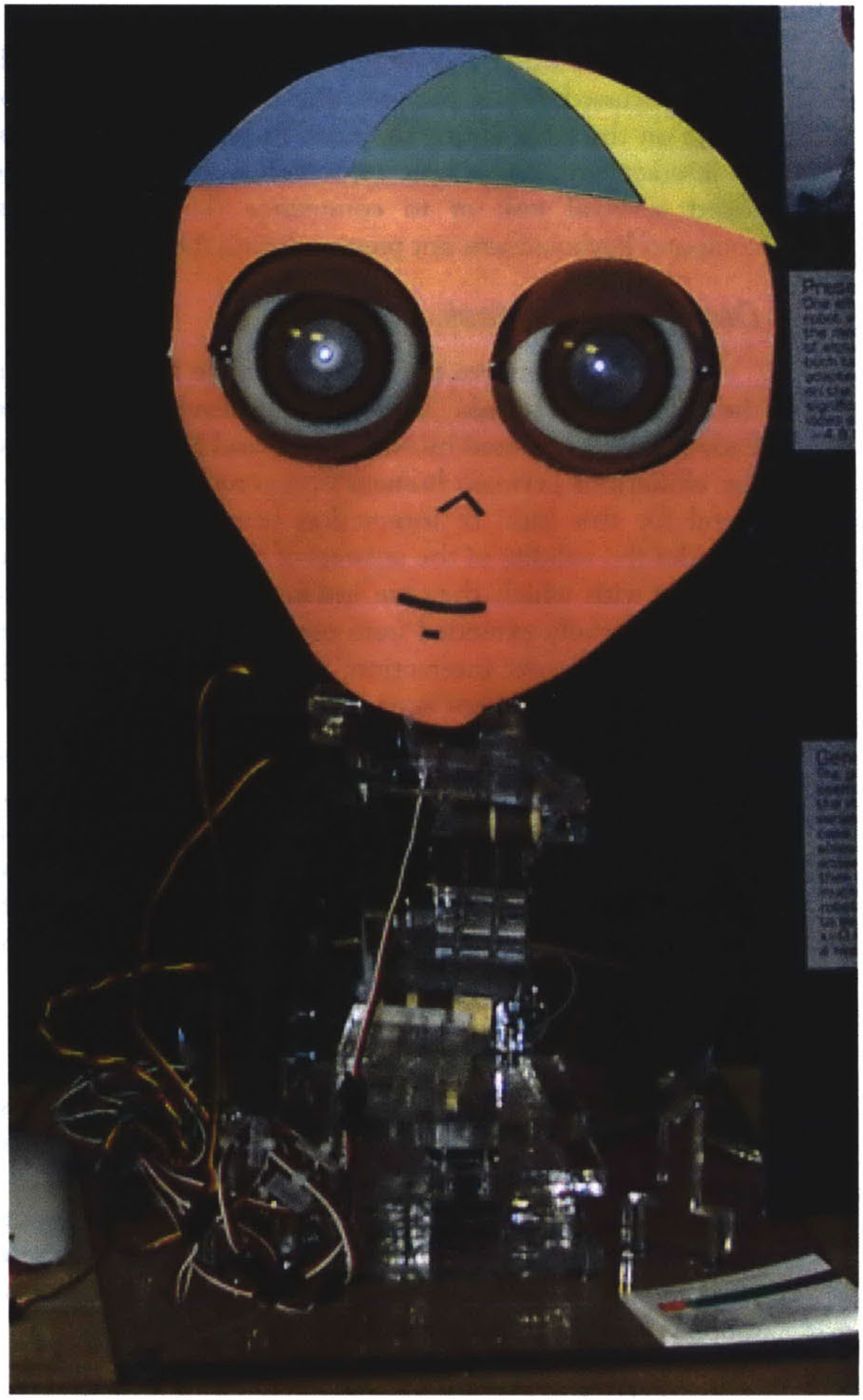

Figure 10. Robot used in second experiment

The topic of this lesson was chosen to be something that would be simple to understand, yet something for which it was unlikely that most of the participants would know much about. The robot read aloud a three to four minute lesson on the Canadian territory of Nunavut. This lesson consisted of some geography (where the territory is located), some history (how it came about politically) and some facts about the inhabitants. The entire lesson came from lessons intended for grade-school children, so there was no aspect of it that should have been difficult for participants to comprehend.

Voice of character

The previous chapter discusses the challenge of choosing a voice to be used for the robot in this type of experiment. In this experiment, the choice was actually made easier because we were not making comparisons across different modalities of interaction, as was the case in the first experiment. Thus we chose to use a recorded human voice, again female, for ease in understanding of the speech. The use of the same voice in this study that was used in the previous study also makes direct comparisons easier between the results of the two experiments.

Robot

The robot in this experiment was an improvement of and elaboration on the robot used in the previous experiment. As discussed in the previous chapter, the robot in the first experiment consisted of a pair of eyes that could move up and down as well as left and right, giving them a full range of motion. There were also upper and lower eyelids on each eye, allowing a blinking movement. As in the previous experiment, the blinking was programmed to roughly correspond to the frequency of human blinking in a similar situation.

For the second experiment, we wanted to take advantage of the fact that a robot shares the same physical space with the participant in the interaction. In order to do this, it is desirable for the robot to be able to move within that shared space. Therefore a robot was built using the same pair of eyes but with the ability to move towards the participant and the computer screen.

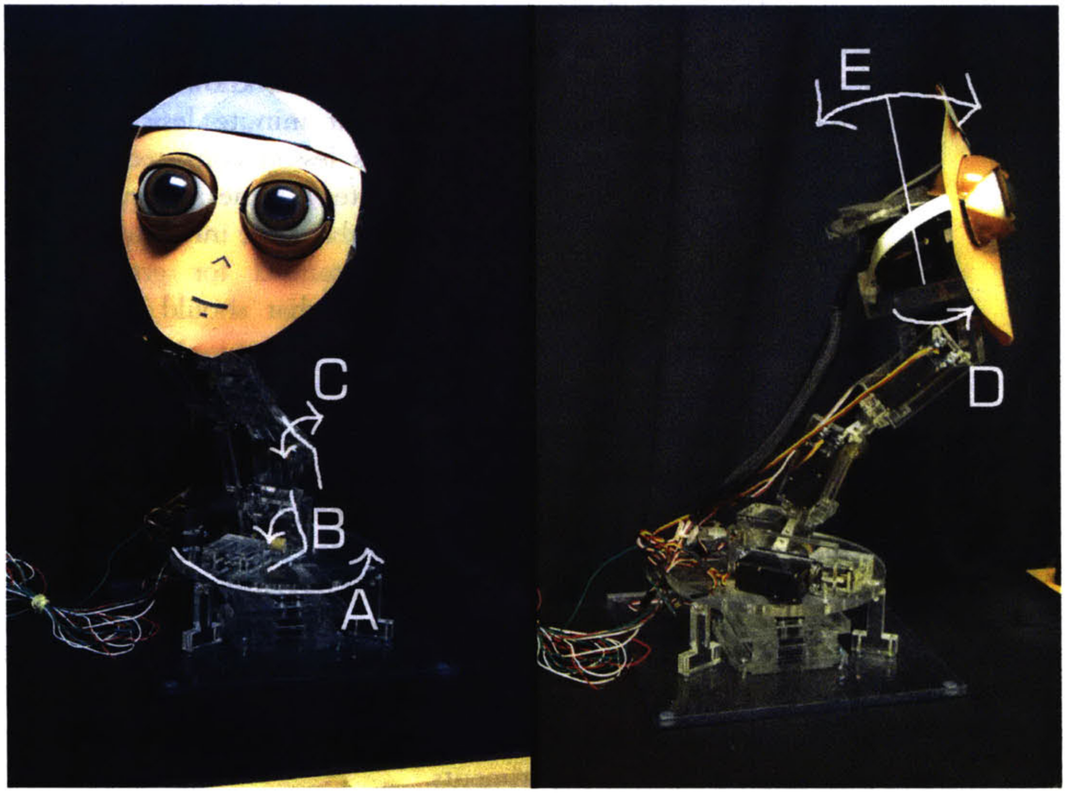

Figure 11. Robot from front and side

To accomplish this, five degrees of freedom were added to the robot. These extra degrees of freedom are shown in Figure 11 and are: (A) the ability to turn at the base, (B) the ability to move forward from the base, (C) forward and backward movement in the middle of the upright section, (D) forward and backward movement at the top, and (E) the ability to turn left and right at the top. This set of joints allowed the robot to move towards the user, the screen, or other objects of interest and still maintain proper orientation of its head. A face was also placed over the robot, giving it a somewhat more anthropomorphic appearance, as can be seen in the images of this section.

This allowed the robot to stand upright, move towards the participant, and move towards the computer screen. In any of these positions, the eyes could be controlled independently, allowing it to look towards the participant, the computer screen, or at another part of the room.

The programming for motions of the robot included a number of pre-scripted animation sequences that it could play out in the proper order during the interactions. One of these actions was for it to move forward from its standing position to look directly at the participant. Another allowed it to turn towards the computer screen and either look at the screen or look back and forth between the participant and the screen. Another action had the robot standing straight up and looking around the room, but not at the participant. Other actions allowed the robot to look at the participant from different points in its range of motion or to look at other fixed points in the room.

The playback of these motions was decided on during the interactions with participants. During the first portion of the desert survival task, when there was no interaction between the participant and the robot, the robot rarely looked toward the participant or the screen, but mostly looked around the room. Once the participant submitted their initial answers, the robot looked toward them to speak to them about the next portion of the experiment. During the interactive period, the robot mostly looked either at the participant or at the computer screen on which the interaction was taking place and only rarely looked away. In the teaching task, the robot mainly looked at the person while either standing straight up or leaning forward and occasionally looked toward the maps that were displayed on the computer screen.

Between-subjects design

The important aspects of the interaction to measure in this experiment concerned differences in participant reactions across two factors: the presence of the robot and the type of task. Because both of the tasks lasted longer than the task in the previous study and were more involved than those tasks, thus requiring greater concentration, we chose a between-subjects design for this experiment. We did not want to introduce the possibility that demand effects on participants might interfere with the results. The data collected was based on questionnaires and behavioral measures that are not known to be difficult to correlate across multiple participants, which allowed this design to be efficacious in this experiment.

4.2.3 Protocol

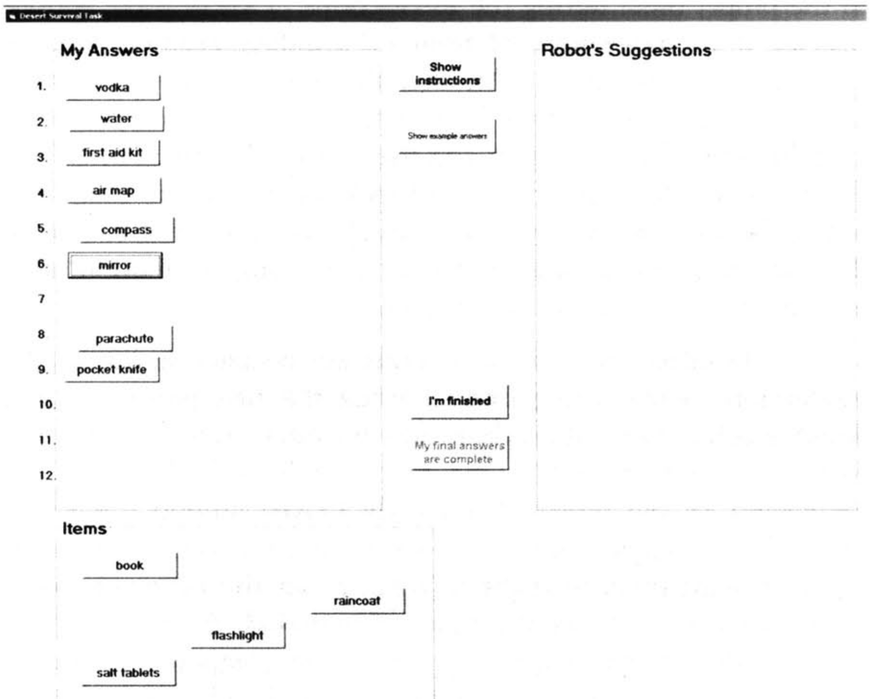

Figure 12. Participant starting Desert Survival Task

Items presented to participants in Desert Survival Task:

first aid kit	book
raincoat	flashlight
vodka	parachute
water	mirror
pocket knife	compass
salt tablets	air map

Figure 13. Desert Survival Task items

Desert Survival Task

Participants were first given instructions about the task that they were about to complete. The experimenter would then start the task on the computer screen and leave the room.

Participants were then given a list of twelve items that may be useful if they were stranded in a desert and asked to rank them in order of importance. See Figure 13 for the list of items given to the participants. Participants were presented the items at the lower left corner of the computer screen on which the interaction took place and were shown a sample of how they could rearrange the items by dragging them to the top of the screen. Look at Figure 12 to see how this would appear to the participant. The item that they placed nearest the top of the screen was the one that they considered most important, the one nearest the bottom was least important, and the rest of the items were ordered somewhere in the middle. Participants did not have to line up the items exactly; they would be interpreted as ordered as long as there was some distinction between their vertical placements inside the response area of the screen. This was done in order to make the task simple for participants to complete. See Figure 14 for an example of what the participant's responses could look like.

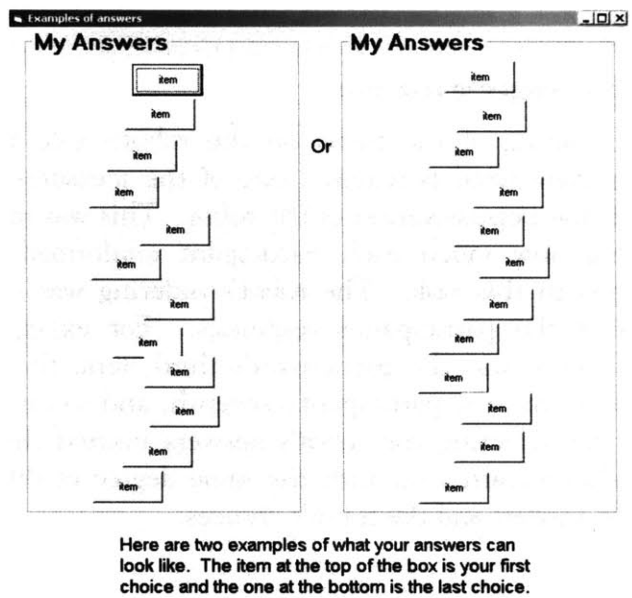

Figure 14. Desert Survival Task sample responses

After deciding on the order in which they valued the twelve items, participants clicked a button on the interface to submit their answers. At this point the software recorded the ordering of the items as the participant's initial answers for this task. The robot would then respond by saying, "Now that you have your first selections made, I can tell you what I think about the items. Just click on any of the items on my side for me to say something about it." At this point, the robot's orderings for the items would appear on the right side of the screen. Figure 15 shows how this appeared to the participant.

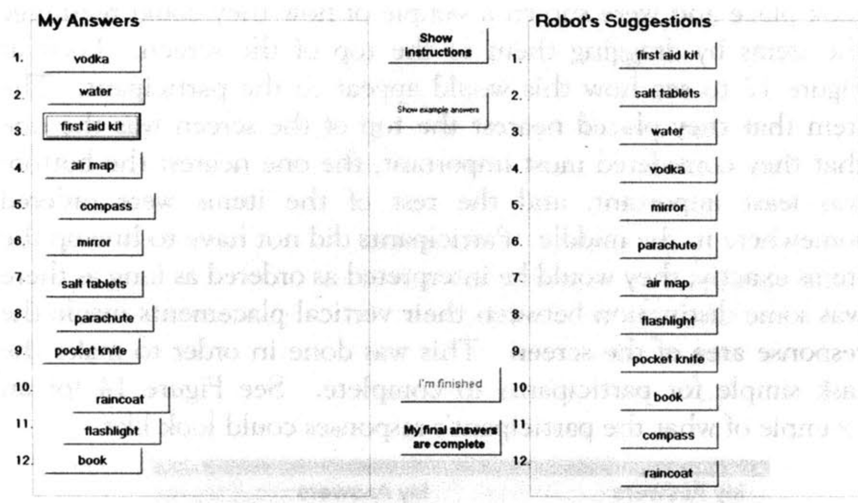

Figure 15. Robot's suggested responses

The ordering of the items on the robot's side was not random. As mentioned previously, one of the measures in this experiment is the persuasiveness of the robot. This was measured by calculating how much each participant conformed to the robot's answers in this task. The robot's ordering was simply a permutation of the participant's responses. For example, the robot's first choice was the participant's third item, the robot's second selection was the participant's seventh, and so on. Using this method for selecting the robot's answers insured that each participant always started off with the same degree of difference between their answers and the robot's choices.

- The first aid kit would be useful in case you get hurt while you are in the desert. It's the only way you could help yourself in that case.
- The pages of the book would be perfect for starting a fire. It can get very cold in the desert at night.
- The raincoat might be useful as protection from the hot sun.
- The flashlight will let you see at night and might help to scare off animals.
- Vodka can be used to sterilize your wounds in case you are injured in the desert.
- The parachute could be spread out on the ground to make your position more visible to search planes.
- Water is extremely important in the desert. You may not find anything else to drink until you get out.
- The mirror can be used for signaling planes or rescuers by reflecting the sun. It's your best chance of being seen at a distance.
- The pocketknife is the only tool that you have available. If you need to make anything out of the other items, this is what you could use.
- The compass will tell you which direction you're heading in if you're trying to walk back to safety.
- The salt tablets will help your body to retain water. Lack of water will be your biggest problem in the desert.
- The map will show you where you need to go if you know where you are.

Figure 16. Robot's statement about each item in desert survival task

The participant could now click on any of the items on the robot's side of the screen to hear the robot give its reasoning for why that item was important. Regardless of the order that the items were listed in, the robot always gave the same response for a particular item. Figure 16 shows the statements made by the robot for each item. During this portion of the interaction, participants could rearrange their answers as they changed their mind about the proper orderings of the items. They were free to listen to the robot's statements about as many or as few of the items as they desired and could listen to the rationale for any item more than once if they chose.

When the participant had decided on a final ordering for the items, they could click a button on the interface to indicate that they were finished. The robot would respond by thanking them, saying "Thanks for working with me today. I hope you enjoyed it." The software running the task would then record the participant's final answers so that the difference between their final and initial answers could be calculated later.

At the conclusion of the interaction, participants were then taken to another room to complete a questionnaire about the robot and the interaction. This questionnaire will be discussed shortly.

Teaching task

In the teaching task, participants were instructed as to the nature of the task and what would be required of them during the interaction. At this point, the experimenter started the task and left the room. A short set of instructions appeared on the computer screen and the task began as soon as the participant clicked the 'Start' button on the screen. (This was the only button that they could click.)

Once the participant initiated the interaction, the robot would begin the lesson. The robot read the lesson aloud to the participant, looking towards the participant most of the time and occasionally moving around as though shifting its position. During the lesson, there were two maps of the geographical region being discussed shown on the computer screen. Participants were informed before the beginning of the lesson that these would be shown throughout and that they were there for reference, but that they were not required, or even able, to do anything with these maps.

After the robot concluded the lesson, it asked the participant a series of nine simple questions about the lesson out loud. Participants were told to answer these questions aloud as best they could. The task was videotaped so that participants' responses could be gathered after the conclusion of the experiment. See Appendix C for the entire lesson that was read to participants and the questions that were asked to them at the end of the lesson. After asking the questions, the robot thanked the participants in the same way as the desert survival task: "Thanks for working with me today. I hope you enjoyed it."

At the conclusion of the interaction, participants were then taken to another room to complete a questionnaire about the robot and the interaction. This questionnaire is discussed shortly.

4.2.4 Dependent measures

There were three types of dependent measures gathered during this experiment. A questionnaire asked about trust, perceived information quality, altruism, engagement, reliability, immediacy, and credibility. Participants' changes in responses during the desert survival task were the second type of measure and were used to judge the persuasiveness of the robot on that task. Another measure of this type is the participants' responses during the teaching task. Finally, behavioral measures of participants' posture and looking patterns towards the robot are used as indicators of the immediacy of the robot.

All subscales of the questionnaire used for this experiment are presented in Appendix B.

4.3 Results

The results of this experiment were analyzed along the scales laid out at the beginning of this chapter. These are trust, perceived information quality, altruism, engagement, reliability, immediacy, credibility, and persuasiveness. Each of these is evaluated along task difference, presence, and gender. These results are presented in the following chapter.

Most analysis was done using a three-way analysis of variance (ANOVA) with the three factors being task (desert survival versus teaching), presence (physically present versus on television screen), and gender of subject. Two exceptions are perceived information quality and immediacy, which were only measured on one of the tasks (desert survival task for perceived information quality and teaching task for immediacy). These measures were analyzed using a two-way ANOVA (presence and gender).

In general, we found that while presence and task do make a difference for some of the dependent variables, this is not the case for all. The gender of participants also turned out to be an important issue in the perception of the robot. Finally, there were a number of interactions between gender and task and between gender and presence that are discussed in the following chapter.

4.3.1 Overview

The data collected in the second experiment was intended to further our understanding of some of the results in the first experiment and to understand new aspects of human-robot interaction. Towards the former aim, the comparison between the interactions in the present and remote robot instances furthers our knowledge of the "real versus fictional" problem that was discussed at the beginning of this chapter. In developing an understanding of new aspects of these interactions, we examine differences caused by the task that a participant is completing with the robot and expanding the number and type of variables that we are interested in measuring.

There were eight things that we were looking for in this experiment. These are trust, perceived information quality, altruism, level of engagement, reliability, immediacy, credibility, and persuasiveness. Each of these items is interesting for reasons discussed in the previous chapter, but all are important to understand for the development of robots that will successfully be able to interact with humans in social situations.

Most of the data that is presented below compares responses or behaviors of one set of participants to another set of participants in a situation where one variable differs. These differences that we are interested in are task (cooperative desert survival task versus teaching task), presence (physically present robot versus robot shown on a television screen), and gender. The statistical analysis that is presented here results from applying the two- and three-way ANOVAs discussed above.

This statistical test is designed to compare samples from groups of participants whose scores are independent of one another [3], as is the case here. The only assumption for this test is that both population samples come from a normal distribution. We have no reason to believe that this assumption does not hold for this data, so we proceed with this test in analysis of the data gathered in this experiment.

In the following sections of this chapter, I report the data by breaking it down into sections along task, presence, gender, and interactions among these three. Each of these sections is further divided into the eight measures that were analyzed.

The data has been grouped into several areas for purposes of discussion. Near the beginning of each of these sections is a graph giving an overview of the data presented there. For more complete information on the data and the statistical tests, please see Appendix D.

4.3.2 Task

In this first section, I present the data analysis broken down by task. In all cases, the variables are presented by comparing the cooperative desert survival task with the teaching task.

What is shown here is that in general, participants ranked the robot higher (or more positively) when they interacted with it on the desert survival task. The only major departure from this is on the scale of credibility, where participants believed the robot was significantly more credible in the teaching task. This seems logical, given the nature of the task. This and other implications of these results are discussed in the following chapter.

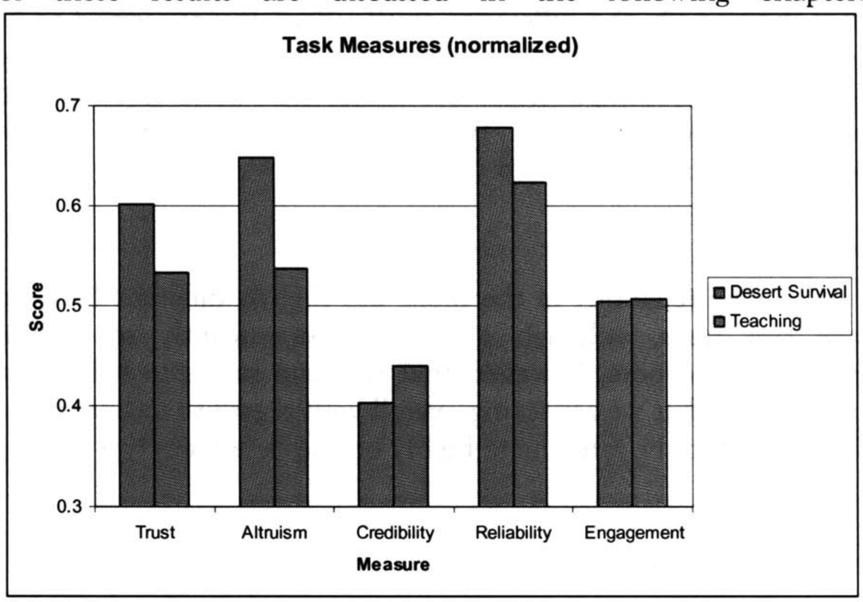

Figure 17. Task measures

Altruism

The degree of altruism that was attributed to the robot was also higher for participants that interacted with it on the more interactive desert survival task. As will be noted shortly, altruism was affected by differences in both task and presence. ($p < 0.01$)

Trust

When comparing levels of trust expressed by participants in the post-experiment questionnaire, we find that they developed slightly greater trust with the robot when cooperating in the desert survival task, rather than completing the teaching task. ($p < 0.05$)

Credibility

As mentioned above, the robot was seen as more credible in the teaching task, in contrast to many of the other results presented in this section, which show a higher outcome to the more interactive desert survival task. The difference here is clearly statistically significant. ($p < 0.05$)

Engagement

The engagement between the participant and the robot did not show much of a difference at all. Participants seemed to be similarly engaged with both characters. The ANOVA results are still given for completeness and shows that there is clearly no difference between these two groups of participants on the question of engagement. ($p < 0.996$)

Reliability

The reliability of the robot was slightly different between the two tasks, although clearly not significantly so. It is interesting to note, however, that the data seems to follow the same trend that we are seeing overall with regard to task: that the robot tends to be rated more highly on the desert survival task. ($p < 0.20$)

4.3.3 Presence

In this section, we are concerned with whether the robot was physically present in front of the participant or was seen on the television screen. We proposed that there would be differences seen between these two situations, which is indeed the case. The strongest difference turns out to be for altruism, although it is interesting to see the lack of difference in some of the other areas. Again, a discussion and analysis of this data will be presented in greater detail in the following chapter.

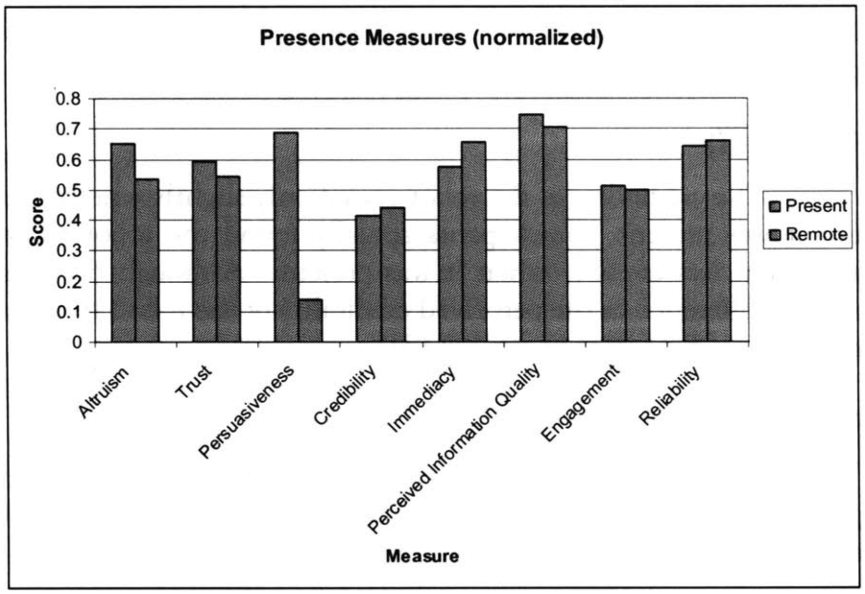

Figure 18. Presence measures

Altruism

Altruism was one measure that was clearly different between the present and remote robots. Participants found the robot that was present to have their interests in mind much more so than for the robot that was not physically in front of them. ($p <$ 0.01)

Trust

As just mentioned, trust is one of the more noteworthy differences regarding the presence of the robot. Although this finding was not significant to the $p < 0.05$ level, it is still interesting to the question of understanding the possible differences among the experimental situations. ($p < 0.07$)

Perceived information quality

Perceived information quality was measured only in the desert survival task, asking about how useful participants felt that the suggestions of the robot were. Therefore all data reported here is from that task. On this measure, participants exhibited a small difference in their beliefs about the information from the robot, finding the information of higher quality when the robot was present. ($p < 0.10$)

Credibility

The credibility of the robot varied only slightly whether or not it was present. Participants actually found the robot to be somewhat less credible when it was present. Although the task made a difference in the perceived credibility of the robot, this was not the case for presence. ($p < 0.13$)

Reliability

Reliability, like engagement, showed little differences for presence or task. There was little difference in this measure across all three variables. ($p < 0.28$)

Immediacy

The degree of immediacy reported by the participant refers to how much it seemed as though the robot was really there and interacting with them. This was not reported above in the task section, as it was not measured in the desert survival task, but only asked about to participants who completed the teaching task. Although it was anticipated that the robot would be considered more immediate when it was present, the data actually shows the opposite trend, although not significantly. ($p < 0.64$)

Engagement

The level of engagement showed almost no difference regardless of whether the robot was present, as can be seen in the data below. In general, it can be seen that engagement did not vary across task or presence. ($p < 0.70$)

Persuasiveness

The persuasiveness of the robot is another measure that was only interesting on one of the tasks, in this case the desert survival task. This was a measure of how much people changed their results to conform to the robot's suggestions during the task. We did find that whether the robot was present made a large difference in the average change of a person's responses, although this change was only nearly significant.

4.3.4 Gender

The final major difference of interest in this experiment was what role the gender of the participant played in the responses to the questions asked and the behaviors exhibited. In this section, the same scales as above are presented, but the data is divided by the participants' gender.

Many of these showed no difference whatsoever in between female and male responses; these will not be shown here. Only those responses that were somewhat different will be shown in this section. Note that this does not mean that gender of the participant did not affect some of the other scales. There are some interactions among participant gender, task, and presence that will be presented in the following section.

Perceived information quality

The reported evaluation of the robot's information varied depending on the gender of the participant. In general women found the information to be more helpful than men, a finding that was predicted based on previous research. ($p < 0.07$)

Reliability

With respect to the participants judging the reliability of the robot, females perceived it to be slightly more reliable, although this difference was definitely not significant. This effect is somewhat more pronounced, however, in the desert survival task and when the robot is present, as seen in the following section on interactions in the data. ($p < 0.17$)

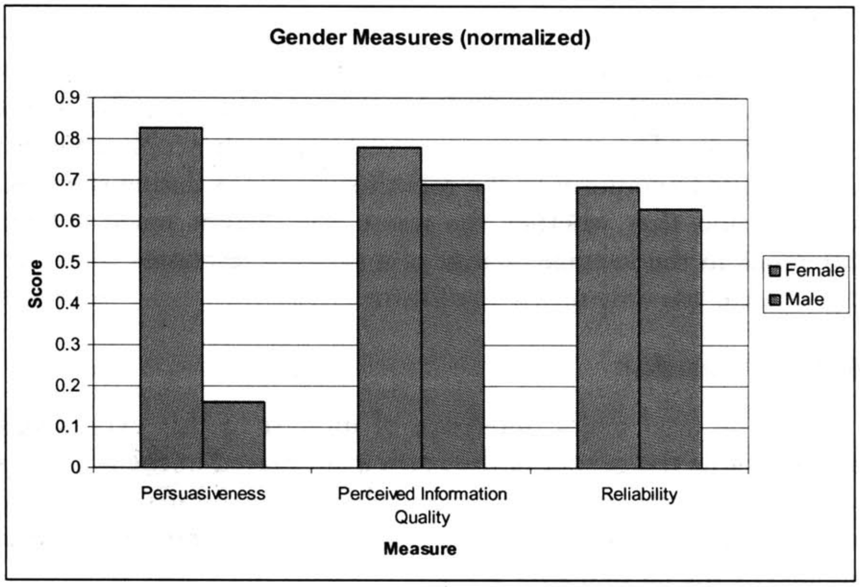

Figure 19. Gender measures

Persuasiveness

The most significant of the effects of gender is seen in the measurements of the persuasiveness of the robot. As can be seen here, the women changed their answers to conform with the robot's suggestions much more so than men.

4.3.5 Interactions Among Task, Presence, and Gender

Finally, we present data from some of the interactions among task, presence, and gender. Although these did not fit into the simpler analysis of the data presented above, there are still several interesting and significant findings that help us compare this study to some of the work that has come before it. Other findings of interest are also presented here.

Gender interactions

Some of the effects of gender are amplified when we look at particular cases instead of averaging over larger sets of data. For example, in the cooperative desert survival task, if we look at only the case when the robot is physically present, we can see that there are several differences among women and men on the scales that we are looking at in this experiment.

The immediacy of the robot shows a presence and gender interaction. In the case where the robot is present, women find it to be significantly less immediate. However men find it to be much more immediate in the present case. ($p < 0.17$)

There is some interaction between task and gender on the trust scale. In the cooperative desert survival task, women find the robot more trustworthy than men do, but the opposite is seen in the teaching task. ($p < 0.11$)

There is also a slight interaction between presence and gender on the trust scale. In this case, women find the physically present robot to be more trustworthy than men do, while men find the remote robot more trustworthy than the women do. (Although it should be noted that both find the physically present robot to be more trustworthy than the remote robot.) ($p < 0.22$)

4.3.6 Distance from robot

One measure of the interactions that we were interested in was how far from the robot or television screen people would sit. The hypothesis was that people would be more interested in the interaction with the physically present robot than the robot depicted on the screen and would therefore sit closer.

The data collected shows that this is indeed the case. As mentioned previously, all participants were videotaped from the front and the side. The side videotape was used to score distance from the robot or screen with the help of a grid that was placed on the wall on the opposite side of the participant as the camera. Analysis was completed with a 2x2 (task, presence) ANOVA and shows that the presence of the robot was clearly significant ($p < 0.001$) and that participants sat closer to the robot than to the screen. We also saw a slight effect of task, in that participants sat closer to the robot during the desert survival task than the teaching task. ($p < 0.16$)

4.3.7 Time looking at robot

A related measure is how long participants spent looking at the robot during the interaction. This comparison was made only between the tasks when the robot was present. (The variance in the television cases was extremely high, with some participants appearing to stare blankly at the screen the entire time and others never looking at it.) We found that in the teaching task, participants looked at the screen a much higher percentage of the time (M = 30.1 % versus M = 6.5%; $t(9) = -3.3$, $p < 0.01$). This is likely due to the nature of the desert survival task, as this task required participants to look at the computer screen most of the time in order to complete the task.

4.3.8 Time of Interactions

Another measure of the interactions was the amount of time that participants spent completing the desert survival task. Because there was no correct answer to the task, this is as close to a measure of successful completion of the task that we have. However, there was little difference in the mean time to completion and a great deal of variance, so we cannot reach any conclusions about the effect of the presence of the robot on this task measure.

4.3.9 Teaching task outcome

The task outcome was easier to measure in the teaching task, where participants answered a series of questions on the content of the lesson at the conclusion of the task. As expected, we found no difference in recall of the information over the short duration of this task.

4.4 Summary

This experiment has allowed us to look in greater detail at the causes of differences among various interactions between humans and robots. The analysis of this data along task, presence, and gender lines tells us more about the causes of the differences that were observed in both this and the previous experiment.

The comparison of this data with that from the first experiment will also elicit more information on the nature of the differences found in engagement in the first experiment. This comparison is made in the next chapter.

In general, we have found that some element of the interaction can alter the perception of any of the variables that were considered in this experiment. In the next chapter, we consider each of these variables in turn, discussing how it was affected by the experimental conditions, and proposing how this may affect future robot designs.

5 Discussion

Now that we have presented the data that was gathered and analyzed in the two experiments, we can discuss how this relates to the research topics that were presented at the beginning of this thesis. We can look at how the different aspects of the interactions that were varied in the experimental designs have an effect on a person's perception of the experience. Finally, we can talk about how this helps us in our quest to create robots that are capable and successful social partners.

5.1 Experimental Questions

The first of the experiments that was carried out looked at the effects of modality on a person's perception of the "other" with which they were interacting. We measured levels of engagement, qualities of the interaction (participants' liking of the character, realness of the character, and ease of interaction), and social presence. In this experiment, we found that a robot was found to be not as easy to interact with nor as engaging as a human. We did show, however, that the robot is easier to interact with and more engaging than the animated character.

This conclusion led directly to the next question: what about the robot makes it different than the animated character or the human? Was it the fact that the robot was a real thing, while the animated character was seen as something fictional that simply appears on the screen, but does not have an existence outside of the screen? Or do these differences come from the fact that the robot is physically present in front of the person while the animated character is seen as remote because it is presented on the screen?

The second study looks at the latter possibility; that the difference results from the physical presence of the robot. In this case, half of our participants interacted with the robot while it was physically present, while the other half saw it on a television. In both cases, the robot responded in the same manner and with the same timings. Even the noise made by the servo motors was audible through the television, leaving the only difference to be the actual physical presence. What we found is that the presence of the robot does indeed have an impact on the assessment made of the robot by the participants. When the robot is physically present, people find it to be more altruistic and more persuasive.

This latter study also explores how the amount of interactivity between a person and a robot affects the person's perceptions of the robot. Here the type of task was varied between solving the highly cooperative and interactive Desert Survival Problem with the robot to listening silently to the robot teaching a lesson on Canadian geography and responding to a short set of questions at the end. What we found here is even more striking than in the case of presence: there is again greater perceived altruism, the information is seen as better, and participants trust the robot more when there is greater interaction. While real-world tasks may not lie at one of these extremes, we can see that there is likely a continuum of interactivity that will affect the way that a robot is perceived and therefore its likelihood of success in a given task.

5.2 Effects of Modality

The main variable in the first experiment and one of the two variables in the second experiment was the modality through which participants interacted with a character. After both experiments, we have looked at interactions between people and another person, an animated character, a physically present robot, and a telepresent robot. What we found is that some of the modality changes make a clear difference on some aspects of presence, level of engagement of the person, and the perceived altruism of the robot.

When we compare the robot to the animated character, we see that two of Lombard and Ditton's conceptualizations of presence were much stronger for the robot. These are presence as transportation and presence as social actor within a medium. This shows that people feel that they are sharing the same space with the robot and that they respond to the social cues of the robot much more so than with the animated character.

What we can draw from this is that there are certain aspects of an interaction that can be easily modified based on the modality of the interaction, while there are other aspects that will vary little or none. What we have shown is that attributions of similarity and identification between the participant and the robot are stronger than that between the participant and the animated character. Social presence is also higher in interactions with the robot than with the animated character. In general, the robot is more well-liked and engaging than the animated character in the task from the first experiment.

5.3 Effects of Proximity

The proximity of the person to the robot was a factor that was considered in the second study. While half of the participants interacted with a robot that was immediately in front of them, the other half operated at a greater perceived distance to the robot, as it appeared to them on the television screen. We found that this had some similar effects to the different modalities that were considered in the first experiment. These effects were not all the same, however, which is why we discuss this separately from the modalities addressed in the previous section.

In the second study, however, the effects were not as strong between the physically present and the remote robots. While we were measuring some different aspects of the interaction than what we measured in the first experiment, we nonetheless expected to find stronger effects. Instead, what we found is that the only significant difference between the two was for the perceived altruism of the character, where the physically present robot was perceived as more altruistic than the remote robot. We did see differences for immediacy, credibility, and persuasiveness, but none of these were at a significant level. One thing that is interesting to note is that there were only slight differences between the present and remote robots for perceived information quality, engagement, and reliability.

These findings lend credibility to the argument that the difference between the robot and the animated character in the first experiment are more a result of the "real versus fictional" rather than the "present versus remote" dichotomy. In other words, I believe that the reason that robots were perceived differently in the first experiment is that they are seen as a real entity that exists in the physical world (although exactly *where* in the physical world is not the most important thing) while the animated character is viewed as a fictional, ephemeral character that appears only for the duration of the interaction.

5.4 Effects of Interactivity and Task Type

Between the two experiments, there were a total of three tasks that were completed by participants in interaction with a robot: the block moving task from the first experiment and the cooperative desert survival task and the teaching task from the second experiment. We compared these three experiments across several measures, mostly concentrating on the tasks from the second experiment, as other factors were controlled for to make these conditions as identical as possible. What we found is that the interactivity in the task is important!

Two of the clearest measures are the level of trust that the participants had in the robot and their perception of the robot's altruism towards them. For both of these measures, the greater the level of interactivity with the robot, the higher they ranked the robot. What seems to be happening is that more interactivity builds trust between the person and the robot and also gives the person a higher opinion of the robot's motives, although further studies will have to be conducted to test this statement.

Following this line of reasoning, we also see a difference in credibility and reliability of the robot, although these were not as significant. (As noted in the next section, both of these are much more significant when we look only at responses from women and not those from men.) Again, this reinforces the belief that the amount of interactivity is an extremely important factor in shaping opinions about the capabilities and qualities of a robot.

One item of note is that we again measured engagement in this study. What we found is that there is no difference between the two tasks. Regardless of the amount of interactivity between the person and the robot, there was no difference in level of engagement. This finding runs counter to our expectations, but is very clearly seen to be the case.

5.4.1 Task outcomes

In order to better understand the comparisons among these three tasks, it is useful to examine some features of the tasks. The three tasks used in these experiments varied in their level of interactivity. We can look at this in two ways: general interactivity (any turn-taking in task and conversation, or manipulation of physical or virtual objects by the participant) and physical interactivity (limited only to manipulation of physical objects by the participant.) We can place the three tasks along the scales created by these definitions. In terms of general interactivity, we find that the teaching task was the least interactive, the block moving task was somewhat more interactive, and the desert survival task involved the greatest level of interaction between the person and the robot or character in the interaction.

One potential advantage of interacting with robots that was posited at the beginning of this thesis is that of it sharing the same physical space with the person or people with whom it is interacting. Thus we should examine how and if the amount of physical interaction in the tasks changed the task or perceptual outcomes. If we turn to a measure of physical interactivity, it is clear that both the teaching task and the desert survival task were lacking in any physical component to the interaction. We may argue that the desert survival task was closer to being interactive, as participants manipulated a computer mouse, but this still does not rise to the level of interaction that is possible, and indeed most interesting, in human-robot interaction. Only the block moving task from the first experiment actually had people manipulating objects that were in the shared space between the person and the robot.

As noted in the chapters describing the experiments, there was very little effect on task outcome, even though there were clear differences on some of the measures that we would expect to affect the task outcome. Two reasons for this may be the short duration of the experiment and the experimental conditions. The duration may have contributed to this similarity because participants did not have enough interaction with the robot (or the other characters) for this to make a difference. It would be desirable to set up an interaction with a longer duration to test the likelihood of a task improvement under what we foresee as more normal conditions for human-robot interaction.

The other possibility for a lack of task outcome differences is the experimental setting under which the interaction took place. Participants may not have performed as they might in solving problems or learning new information in their daily lives. It will take more work before we are ready to deploy robots in everyday environments outside of the lab to determine whether this is a factor in our results.

5.5 Effects of Gender

The final major area of difference that we found in this work was the effect of gender of the participant on the perceived qualities of the interaction. While this was not surprising based on earlier work that was discussed in the background chapter, it is nonetheless interesting to see where our work agrees with and potentially refutes earlier findings.

Both credibility and reliability are significantly higher in the more interactive task for women, although this effect is not as strong for men. In the highly interactive task, women also perceived the quality of the information provided by the robot to be higher than men did. Women also were more persuaded by the robot's suggestions in the Desert Survival Task, which was seen through their modifying their answers to conform to the answers suggested by the robot more so than men did.

One gender difference that was noted in the background was that Reeves et al. found that women found their robot to be more credible than men after three short interactions [43]. In the interactive task in the second study presented here, credibility of the robot was measured and shows no difference between women and men. (Women $\underline{M}$=44.73, Men $\underline{M}$ =44.71) The nature of the tasks may create some of the difference between the two findings, but probably not all. There are likely other aspects of the interactions that appear differently to women and men that were not explored.

The question remains: why do we see differences between men and women in their reactions to our robot? There are several possible explanations that I will put forth here, although it is clear that further work will be needed in order to answer this question. One explanation is that the robot was inadvertently constructed to appeal more to women. This may have lead to their inclination to find it more reliable, more persuasive, and to provide higher quality information when compared to men's responses. It is possible that the look of the robot is very important in determining these responses. If this is the case, we clearly need to understand more about this, as many of the applications for robots in the future may well depend on their ability to provide information to a person and that information must be believed to make these types of interactions a success.

Another possibility is that it was the voice, and not the appearance, of the robot that swayed the responses. In all cases where the robots (or other characters) spoke in the studies presented in this thesis, they did so with a recorded female voice. Much as the appearance may change responses, it is possible that voice could have that effect. Again, further study is needed to see if this could be the case, as it will impact future communication strategies when constructing robots for human interaction.

5.6 What Does This Mean?

One of the most important findings here is the difference between the robot and the animated character in the first study. As the second experiment brought to light, this appears to be because of the real, physical existence of the robot. When this is the case, people are more engaged and feel a greater sense of social presence. This affirms that idea that we started with: robots are indeed a good interaction partner and provide definite benefits over interactions with an animated character.

When we turn to the differences among interactions between people and robots, we also have some lessons to draw upon for constructing future interactions. Notably, the amount of interaction between a person and a robot plays an important part in influencing the perception of that robot. A number of effects on a person's beliefs about a robot have been shown to be changed by the amount of interaction that a person has with a robot. Smaller effects were shown for the physical presence of the robot as well, but these do not seem to be as important in shaping the interaction.

Finally, the effects of gender are of importance when designing a robot. Further work should be done in exploring why the gender differences that were shown here exist. Maybe we could design the robot so that it could engender greater trust or exert more influence over men in some cases. Or create a robot in such a way that women are not as persuaded by its suggestions.

Findings summarized:

Presence matters...

When a robot is physically present:

- people trust it more,
- find it more altruistic,
- but are no more engaged than if it is remote.

...but Interaction matters more

When there is a greater amount of interactivity between a robot and person:

- people find it more altruistic,
- find the information to be of higher quality (for women only),
- find it more persuasive,
- trust it more,
- but are still not significantly more engaged than with less interaction.

And the Gender of the participant is important

Women find the information from the robot to be of higher quality than men do.

Women believe that the robot is more reliable than men do.

Women are more likely to be persuaded by the information coming from the robot than men are.

Figure 20. Summary of research findings

5.6.1 Agreement with Previous Research

The first figure in the Background chapter (Figure 1) listed ten of the most relevant findings that are discussed in that chapter. Figure 20 shows the main findings of the work from this thesis. We can also address how this work agrees or disagrees with each of the points made in the earlier summary as well.

The first section of Figure 1 addressed the information factors of credibility and information quality. We can affirm that presence does have some impact on credibility and perceived information quality, although our results were not statistically significant. The second point had to do with the distance between task partners. This same finding would indicate that the correlation between distance and credibility is the case, but we did not vary the supposed difference between the robot and person enough to draw any conclusions on this point. Finally, previous work showed that a closer affinity between a person and robot led to higher evaluations of quality. We did not measure affinity, but we did look at engagement between the person and robot and showed only a weak correlation between these two measures (r = 0.36).

The second section of Figure 1 discusses task performance and pointed out that others have found no relation between task performance and proximity. This holds in our task measures across both tasks of the second experiment. We found little difference in task outcome regardless of the presence of the robot. The second point in this section is that the presence of the robot makes a difference in men's memory capabilities, while making no difference for women. As we did not test recall, we can not make a comparison with this finding.

The next section addresses the likeability of a robot. Earlier work has shown that a physically present robot will be more liked than a remote one and that a robot that uses humor will be more liked than a serious (or dull) robot. Humor was not measured in our studies, nor was likeability in our second study.

Two findings were reported on social judgments made by people about their interaction partners. One specifically concerned robots, noting that women see them as more of a social equal than men do. The other is that people make higher social judgments about a partner who is physically closer. Regarding the first point about social equals, we asked participants to rate the robot on a 7-point scale as to whether they agree with the statement "He/she is a lot like me." We found no significant difference for women and men on this scale, leading us to question whether this difference generally exists. On the second point, we again did not vary the perceived distances enough to be able to make a judgment. We do note, however, that when the robot was physically present instead of seen only on the television screen, people found it to be more trustworthy and altruistic, which lends support to this theory.

The final section noted that greater levels of interaction occur when partners are physically closer to one another. Because we did not vary distance enough, it is difficult to draw conclusions. We can say that the presence of the robot (not unlike the use of distance in the study by Burgoon, et al. cited in the second chapter [13]) had very little effect on the level of engagement of the person with the robot, as shown in the results of the second study. As this previous finding came from work on human interactions, we should do further research to determine if indeed this principle does not hold for human-robot interaction.

5.7 Summary

At the beginning of this thesis, there was an extensive discussion of experiments that are related to the work that was performed in this set of studies. We also discussed several hypotheses that are reported on here.

We said that robots that are perceived as closer will be seen as more credible, more persuasive, and will score higher on judgments of a social nature. Indeed, we found that the present robot is seen as more credible, but not significantly so (t test yields $p < 0.20$). The robot was also more persuasive when present. (For persuasiveness, t test gives $p < 0.14$.) As for social judgments, we found that the present robot rated higher on scales of altruism, and immediacy than the remote robot.

We also predicted that a robot will be found more engaging than an animated character. As noted above, this was the case. In the first study, we found higher levels of engagement when the participants were interacting with the robotic character than when they were interacting with the animated character.

Another hypothesis was that greater interactivity between person and robot would yield greater trust and evaluation of the performance of the robot. In the Effects of Interactivity and Task Type section above, the relevant findings which show this to be the case were discussed.

6 Conclusion

6.1 Applicability to Particular Domains

In the first chapter of this thesis, a variety of uses for robots were introduced as the motivation for this work. I would now like to revisit some of those potential applications and address how this work impacts the building of these kinds of robots.

The first area that was addressed was that of robots in entertainment. This encompasses a number of different ideas, from robots as actors in movies to robotic pets and dolls for children. This is one of several areas where the interaction results that were found in this work could be important. Greater levels of interaction between the robot and person lead to higher opinions of the robots intentions and greater trust in the robot, both of which might be desired in a toy. This must be tempered with the finding that people also find the robot more persuasive and its information more believable. Depending on the design of the robotic character, these could be positive or negative aspects to the interaction.

One important use of robots for which there is currently a project under development is the use of robots as partners in scientific exploration. The example that was given in the first chapter, Robonaut, is the project that is currently being developed. In this capacity, a robot not only needs to be trustworthy and reliable, but must be seen that way by the humans who are interacting with it in order to make these kinds of interactions successful. For these reasons, the fact that there is a physical robot that can interact with people is important. The type of work will lend itself well to a lot of interaction between the robot and its human partners, which will work well to increase the trust in the robot as long as it is successful in performing its tasks. One finding from this work that is worth trying to understand more about is why women find a robot more reliable than men. While this may be because of the short nature of the interactions in these experiments, knowing the answer to this may help to improve the relationship between humans and scientific robots.

In terms of applicability of the findings of this work, we can group together several of the other types of robots that are envisioned for the future. There will be some shared design goals when constructing robots for communication, education, or information. In order for any of these robots to achieve successful interactions, they must be trusted and be engaging. Both presence and level of interaction impact both of these qualities of a robot.

6.2 Robots as Partners

One foundation upon which this work was built is the idea of robots as capable partners in interactions with humans. This ranges from scientific exploration to teaching and communication. The fact that we found significant results in the two experiments presented in this thesis reinforces our belief in the validity of this concept. Much as Reeves and Nass reported in *The Media Equation*, we have seen that there are aspects of the social elements of human-robot interaction that can be influenced by the mere presence of the robot or the way that it acts with people. This presents us with the likelihood of being able to shape these interactions in part with the way that a robot behaves.

Another benefit of this work is that we have delineated specific aspects of interactions that can be influenced by the presence of a robot, by the level of interactivity between a robot and person, and by the gender of the person in a given interaction. These influences are reported in the previous chapters and give us insight into how to build robots to accomplish specific goals, as noted in the preceding section.

6.3 Future Work

Although we noted some effects in the previous chapter on the differences caused by the amount of physical interaction between a person and a robot, there is still a lot of work to be done in this area. Open questions include understanding how the amount of interactivity affects perceptions of people and task outcomes, whether the robot must participate in the manipulation (as in Robonaut) or it can simply direct actions while being aware of the physical space (as in our first experiment where participants moved blocks at the request of the robot), and how qualities of the movement (smooth, lifelike motion versus jerky, "robotic" movement or fast versus slow motion) affect a person's perception of the robot.

As discussed at the conclusion of the last chapter, more work remains to be done on understanding how the level of interactivity will affect the perceptions of a robot. The intuition developed through carrying out this work leads to the idea that more interactivity is better. However, this may be moderated by the type of interaction; in some cases more interactivity could interfere with the task. There is also necessarily a limit on the amount of interaction that is desirable. Up to some point the general rule will hold; after that, the increase in interaction will detract from the goals of the task at hand.

The next major step in this work is to construct a robotic system that can interact with people on an extended basis to carry out one of the types of interactions that have been discussed throughout this thesis. This step will validate the findings presented here and help us to move one step closer to realizing our vision of sociable robots that can assist people in their everyday lives.

References

1. Allen, K., Blascovich, J., Tomaka, J. and Kelsey, R. Presence of Human Friends and Pet Dogs as Moderators of Autonomic Responses to Stress in Women. *Journal of Personality and Social Psychology*, *61* (4). 582-589, 1991.
2. Argyle, M. and Cook, M. *Gaze and mutual gaze.* Cambridge University Press, Cambridge, 1976.
3. Aron, A. and Aron, E.N. *Statistics for Psychology.* Prentice-Hall, Upper Saddle River, New Jersey, 1999.
4. Baltus, G., Fox, D., Gemperle, F., Goetz, J., Hirsch, T., Magaritis, D., Montemerlo, M., Pineau, J., Roy, N., Schulte, J. and Thrun, S., Towards Personal Service Robots for the Elderly. in *Proceedings of the Workshop on Interactive Robotics and Entertainment (WIRE)*, (Pittsburgh, PA, 2000), Computer Science and Robotics, Carnegie Mellon University, 7.
5. Bengtsson, B., Burgoon, J.K., Cederberg, C., Bonito, J.A. and Lundeberg, M. The Impact of Anthropomorphic Interfaces on Influence, Understanding, and Credibility.
6. Bickmore, T.W. Relational Agents: Effecting Change through Human-Computer Relationships *Media Arts and Sciences*, Massachusetts Institute of Technology, Cambridge, MA, 2003, 284.
7. Bradner, E. and Mark, G., Why Distance Matters: Effects on Cooperation, Persuasion and Deception. in *CSCW '02*, (New Orleans, LA, 2002), ACM, 226-235.
8. Breazeal, C. Designing Sociable Machines: Lessons Learned. in Dautenhahn, K., Bond, A.H., Canamero, L. and Edmonds, B. eds. *Socially Intelligent Agents: Creating Relationships with Computers and Robots*, Kluwer Academic Publishers, Norwell, MA, 2002, 149-156.
9. Breazeal, C. Toward sociable robots. *Robotics and Autonomous Systems*, *42*. 167-175, 2003.
10. Breazeal, C.L. *Designing Sociable Robots*. The MIT Press, Cambridge, MA, 2002.
11. Breazeal, C.L. Sociable Machines: Expressive Social Exchange Between Humans and Robots *EECS*, MIT, Boston, 2000, 264.
12. Burgoon, J.K., Bonito, J.A., Bengtsson, B., Cederberg, C., Lundeberg, M. and Allspach, L. Interactivity in human-

computer interaction: a study of credibility, understanding, and influence. *Computers in Human Behavior*, *16*. 553-574, 2000.

13. Burgoon, J.K., Bonito, J.A., Ramirez, A., Jr., Dunbar, N.E., Kam, K. and Fischer, J. Testing the Interactivity Principle: Effects of Mediation, Propinquity, and Verbal and Nonverbal Modalities in Interpersonal Interaction. *Journal of Communication*. 657-677, 2002.
14. Camarinha-Matosa, L.M. and Vieira, W. Intelligent mobile agents in elderly care. *Robotics and Autonomous Systems*, *27* (1-2). 59-75, 1999.
15. Cary, M.S. The Role of Gaze in the Initiation of Conversation. *Social Psychology*, *41* (3). 269-271, 1978.
16. Casper, J. Human-Robot Interactions During the Robot_Assisted Urban Search and Rescue Response at the World Trade Center *College of Engineering*, University of South Florida, 2002, 114.
17. Corporation, O. OMRON NeCoRo Cat-Like Robot, Omron Corporation, 2003.
18. Dahlback, N., Jonsson, A. and Ahrenberg, L., Wizard of Oz Studies - Why and How. in *Intelligent User Interfaces '93*, (1993), ACM, 193-200.
19. Dawson, M.E., Schell, A.M. and Filion, D.L. The electrodermal system. in Cacioppo, J.T. and Tassinary, L.G. eds. *Principles of psychophysiology: Physical, social, and inferential elements*, Cambridge University Press, Cambridge, 1990, 295-324.
20. Fong, T., Nourbakhsh, I. and Dautenhahn, K. A Survey of Social Robots, Carnegie Mellon University Robotics Institute, Pittsburgh, PA, USA, 2002, 56.
21. Goetz, J. and Kiesler, S., Cooperation with a Robotic Assistant. in *Conference on Human Factors In Computing Systems (CHI 2002)*, (Minneapolis, MN, USA, 2002), 578-579.
22. iRobot. Official Roomba Home Page, 2003.
23. John, O.P., Donahue, E.M. and Kentle, R. The Big-Five Inventory -- Versions 4a and 54, Institute of Personality and Social Research, University of California, Berkeley, CA, 1991.
24. Kahn Jr., P.H., Friedman, B. and Hagman, J., "I Care About Him as a Pal": Conceptions of Robotic Pets in

Online AIBO Discussion Forums. in *CHI 2002*, (Minneapolis, MN, 2002), ACM, 632-633.
25. Kiesler, S. and Goetz, J., Mental Models of Robotic Assistants. in *Conference on Human Factors In Computing Systems (CHI 2002)*, (Minneapolis, MN, USA, 2002), 576-577.
26. Koda, T. and Maes, P., Agents with Faces: The Effect of Personification. in *Fifth IEEE International Workshop on Robot and Human Communication*, (1996), IEEE, 189-194.
27. Lafferty, J.C., Eady, P.M. and Elmers, J. The Desert Survival Problem. in *Experimental Learning Methods*, Plymouth, Michigan, 1974.
28. Latane, B. The Psychology of Social Impact. *American Psychologist, 36* (4). 343-356, 1981.
29. Latane, B., Liu, J.H., Nowak, A., Bonevento, M. and Zheng, L. Distance Matters: Physical Space and Social Impact. *Personality and Social Psychology Bulletin, 21* (8). 795-805, 1995.
30. Lombard, M. and Ditton, T.B. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3* (2). University of Southern California Annenberg School for Communication, 1997.
31. Lombard, M., Ditton, T.B., Crane, D., Davis, B., Gil-Egul, G., Horvath, K. and Rossman, J., Measuring Presence: A Literature-Based Approach to the Development of a Standardized Paper-and-Pencil Instrument. in *Presence 2000: The Third International Workshop on Presence*, (Delft, The Netherlands, 2000).
32. Moon, Y., The Effects of Distance in Local versus Remote Human-Computer Interaction. in *CHI 98*, (Los Angeles, CA, 1998), ACM, 103-108.
33. Morkes, J., Kernal, H.K. and Nass, C., Humor in Task-Oriented Computer-Mediated Communication and Human-Computer Interaction. in *CHI 98*, (Los Angeles, CA, 1998), ACM, 215-216.
34. NASA. Instructing Robotic Assistants to Acquire Autonomous Behaviors; Acquisition of Autonomous Behaviors for Robotic Assistant Through Instruction, NASA Johnson Space Center, Houston, TX, USA, 2002, 37.

35. NASA. Robonaut, National Aeronautics and Space Administration, 2001.
36. Nass, C. Stanford meeting. Kidd, C. ed., Stanford, CA, 2002.
37. Nass, C., Fogg, B. and Moon, Y. How Powerful Is Social Identity? Affiliation Effects in Human-Computer Interaction, Stanford University Communications Department, Stanford, 1995.
38. Nass, C., Steuer, J. and Tauber, E., Computers are Social Actors. in *CHI '94*, (1994).
39. Nass, C., Steuer, J., Tauber, E. and Reeder, H., Anthropomorphism, agency, and ethopoeia: computers as social actors. in, (1993), ACM Press.
40. Pertemps.co.uk. Office Staff Get Close to Violence, Pertemps, 2002.
41. Raveis, V.H., Mesagno, F., Darus, D. and Gottfried, D. Psychological Consequences of Pet Ownership of Terminally Ill Cancer Patients, Memorial Sloan-Kettering Cancer Institute, Department of Social Work Research, New York, New York, 1993.
42. Reeves, B. and Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, Cambridge, England, 1996.
43. Reeves, B., Wise, K., Maldonado, H., Kogure, K., Shinozawa, K. and Naya, F., Robots Versus On-Screen Agents: Effects on Social and Emotional Responses. in *CHI 2003*, (Fort Lauderdale, 2003), ACM, (in press).
44. Rubin, R.B., Palmgreen, P. and Sypher, H.E. (eds.). *Communication Research Measures: A Sourcebook*. The Guilford Press, New York, 1994.
45. Scassellati, B. How Robotics and Developmental Psychology Complement Each Other, Michigan State University, Lansing, MI, 2000, 7.
46. Surgical, I. Intuitive Surgical da Vinci robotic surgical system, Intuitive Surgical, 2003.
47. Turkle, S. and Daste, O. Aibo and My Real Baby Interactions with Elderly People, Boston, MA, USA, 2003.
48. UN, U.N.P.D. World Population Ageing: 1950-2050, United Nations, New York, 2002.

49. Vilhjalmsson, H.H. Avatar Augmented Online Conversation *Media Arts and Sciences*, Massachusetts Institute of Technology, Cambridge, MA, 2003, 162.
50. Winston, S. Personal communication, 2002.

Appendix A – Experiment I Questionnaire

Interactive Experiences Questionnaire

Thank you very much for completing the interaction with the characters and for agreeing to complete this questionnaire.

Filling out this questionnaire is completely voluntary. You may choose to not answer any or all of the questions. Any answers you do provide will be kept strictly confidential and used only in evaluating our research.

There are four parts to this form, asking about the interactions that you just experienced and some general information about yourself. Completing the entire questionnaire should take you approximately 15 minutes.

Instructions:

The questions on these pages ask about the interactive experience you just had.

There are no right or wrong answers; please simply give your first impressions and answer all of the questions as accurately as possible, even questions that may seem unusual or to not apply.

Many of the questions will have space for you to respond about all three of the characters: the robotic eyes, the animated character on the screen, and the human eyes. Please think about your experiences with each of the three when answering the questions.

Please circle the responses that best represent your answers. All of your responses will be kept strictly confidential.

If you have any questions at any time while you are completing this questionnaire, please ask the experimenter for further explanation.

Part I. General impressions (2 pages)

Please circle the number that best represents your answer about each character.

How engaging was the interaction?

Robot:	Not at all	1	2	3	4	5	6	7	Very much
Screen:	Not at all	1	2	3	4	5	6	7	Very much
Human:	Not at all	1	2	3	4	5	6	7	Very much

How relaxing or exciting was the experience?

Robot:	Very relaxing	1	2	3	4	5	6	7	Very exciting
Screen:	Very relaxing	1	2	3	4	5	6	7	Very exciting
Human:	Very relaxing	1	2	3	4	5	6	7	Very exciting

To what extent did you experience a sensation of reality?

Robot:	Not at all	1	2	3	4	5	6	7	Very much
Screen:	Not at all	1	2	3	4	5	6	7	Very much
Human:	Not at all	1	2	3	4	5	6	7	Very much

How much attention did you pay to the display devices or equipment rather than to the interaction?

Robot:	Not at all	1	2	3	4	5	6	7	Very much
Screen:	Not at all	1	2	3	4	5	6	7	Very much
Human:	Not at all	1	2	3	4	5	6	7	Very much

How often did you feel that the character was really alive and interacting with you?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

How completely were your senses engaged?

Robot:	Not at all	1	2	3	4	5	6	7	Very much
Screen:	Not at all	1	2	3	4	5	6	7	Very much
Human:	Not at all	1	2	3	4	5	6	7	Very much

How well were you able to view the character from different angles?

Robot:	Not well at all	1	2	3	4	5	6	7	Very well
Screen:	Not well at all	1	2	3	4	5	6	7	Very well
Human:	Not well at all	1	2	3	4	5	6	7	Very well

How responsive was the environment you saw/heard to any actions that you initiated during the media experience?

Robot:	Not at all	1	2	3	4	5	6	7	Very much
Screen:	Not at all	1	2	3	4	5	6	7	Very much
Human:	Not at all	1	2	3	4	5	6	7	Very much

How natural was the interaction with the character?

Robot:	Not at all	1	2	3	4	5	6	7	Very much
Screen:	Not at all	1	2	3	4	5	6	7	Very much
Human:	Not at all	1	2	3	4	5	6	7	Very much

It seemed like the events I saw/heard were actually occurring at the time I saw/heard them instead of being recorded and replayed.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

It seemed that the events I saw/heard had been recorded at an earlier time and were being replayed.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

It seemed that the events I saw/heard had happened at an earlier time and were being replayed out of order - they were edited together later.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

The experience caused real feelings and emotions for me.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I was so involved in the interaction that I lost track of time.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

After the interaction ended I had to adjust back to the immediate physical surroundings.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

Part II. Characteristics of the three interactions (5 pages)

For each of the pairs of words below, please circle the number that best describes your evaluation of the media experience.

With the character played by the ROBOT eyes

Impersonal	1	2	3	4	5	6	7	Personal
Unsociable	1	2	3	4	5	6	7	Sociable
Insensitive	1	2	3	4	5	6	7	Sensitive
Dead	1	2	3	4	5	6	7	Lively
Unresponsive	1	2	3	4	5	6	7	Responsive
Unemotional	1	2	3	4	5	6	7	Emotional
Remote	1	2	3	4	5	6	7	Immediate

With the character played by the eyes on the SCREEN

Impersonal	1	2	3	4	5	6	7	Personal
Unsociable	1	2	3	4	5	6	7	Sociable
Insensitive	1	2	3	4	5	6	7	Sensitive
Dead	1	2	3	4	5	6	7	Lively
Unresponsive	1	2	3	4	5	6	7	Responsive
Unemotional	1	2	3	4	5	6	7	Emotional
Remote	1	2	3	4	5	6	7	Immediate

With the character played by the HUMAN eyes

Impersonal	1	2	3	4	5	6	7	Personal
Unsociable	1	2	3	4	5	6	7	Sociable
Insensitive	1	2	3	4	5	6	7	Sensitive
Dead	1	2	3	4	5	6	7	Lively
Unresponsive	1	2	3	4	5	6	7	Responsive
Unemotional	1	2	3	4	5	6	7	Emotional
Remote	1	2	3	4	5	6	7	Immediate

How often did you want to or did you move your body or part of your body either closer to or further away from the characters you saw/heard?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

How often did you want to or did you make eye contact with the character?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

To what extent did you feel you could interact with the character?

Robot:	None	1	2	3	4	5	6	7	Very much
Screen:	None	1	2	3	4	5	6	7	Very much
Human:	None	1	2	3	4	5	6	7	Very much

How often did you have the sensation that the character could also see/hear you?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

How much control over the interaction with the character did you feel that you had?

Robot:	None	1	2	3	4	5	6	7	Very much
Screen:	None	1	2	3	4	5	6	7	Very much
Human:	None	1	2	3	4	5	6	7	Very much

How often did you make a sound out loud (e.g., laugh, speak) in response to someone you saw or heard in the interaction?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

How often did you smile in response to the character?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

How often did you want to or did you speak to the character?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

How often did it feel as if the character was talking directly to you?

Robot:	Never	1	2	3	4	5	6	7	Always
Screen:	Never	1	2	3	4	5	6	7	Always
Human:	Never	1	2	3	4	5	6	7	Always

Give your impression about each statement with regard to the three characters by circling one number for each character.

He/she is a lot like me.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I would like him/her to be a friend of mine.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I would like to talk with him/her.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

If he/she were feeling bad, I'd try to cheer him/her up.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I looked at him/her often.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

He/she seemed to look at me often.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

He/she makes me feel comfortable, as if I am with a friend.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I like hearing his/her voice.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

If there were a story about him/her in a newspaper or magazine, I would read it.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I like him/her.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

I'd like to see/hear him/her again.

Robot:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Screen:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree
Human:	Strongly disagree	1	2	3	4	5	6	7	Strongly agree

For each word, give your overall impressions about each character by circling one number in each column for each characteristic

	Robot							Screen							Human						
	Describes Poorly						Describes Well	Describes Poorly						Describes Well	Describes Poorly						Describes Well
Annoying	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Balanced	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Compelling	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Convincing	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Credible	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Enjoyable	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Entertaining	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Fair	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Favorable	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Good	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Helpful	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Honest	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Homogeneous	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Informative	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Likable	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Negative	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Persuasive	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Reliable	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Satisfying	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Trustworthy	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Useful	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Varied	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Well-composed	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Part III. Overall impressions (1 page)

Which character in the environment you saw/heard did you most want to interact with during the media experience? (Please circle one)

The robot | The animated character on the screen | The human character

Which character was more interested in the task?

The robot | The animated character on the screen | The human character

Which character was more interested in you?

The robot | The animated character on the screen | The human character

What was missing from the animated character on the screen that would make it seem more alive?

What was missing from the robotic character that would make it seem more alive?

Did you enjoy the interaction with any one of them more than the others? If so, why?

Part IV. Biographical information (1 page)

You're almost done! These last questions are about you. Again, all of your responses will be kept strictly confidential, so please answer as accurately and honestly as possible to help us in evaluating the responses.

How old are you (in years)? ______

Please indicate your gender: ____ Male ____ Female

What is your race?
____ Asian
____ African American
____ Hispanic
____ Pacific Islander
____ White
____ Other

What is your level of education?
____ Some high school
____ High school degree
____ Some college
____ College degree
____ Some graduate school
____ Graduate school degree

What is your occupation? ______________________________

How many hours do you spend watching television (including watching videotapes) in a typical day? (estimate as closely as possible)
____ 0 hours
____ Less than 1 hour
____ 1 or 2 hours
____ 3 or 4 hours
____ 5 or 6 hours
____ 7 hours
____ More than 7 hours

How often do you use a video game system (at home, work, school, or at an arcade)?
____ Never
____ Less than once a month
____ 1-4 times a month
____ 5-10 times a month
____ 11-20 times a month
____ More than 20 times a month

Do you own or have you played with a robotic toy? (e.g., Sony AIBO)
____ Never
____ Own one or more
____ Played with them

How many times have you used an interactive virtual reality system
____ Never
____ 1 time
____ 2-4 times
____ 5-7 times
____ 8 or more times

How much do you know about robotics?

None 1 2 3 4 5 6 7 A lot

How much do you know about artificial intelligence?

None 1 2 3 4 5 6 7 A lot

Appendix B – Experiment II Questionnaire

Note that all measures of Cronbach's alpha were calculated based on the data reported for the second experiment.

Subscale: Trust

Source: Receptivity/Trust subscale of Relational Communication Scale found on pages 308-13 of Communication Research Measures [44]

Use: Both tasks

Calculated Cronbach's alpha: 0.79

All items were ranked on a seven-point Likert scale with a range of "strongly disagree" (1) to "strongly agree" (7).

The robot was sincere.

The robot was interested in talking with me.

The robot wanted me to trust him/her.

The robot was willing to listen to me.

The robot was open to my ideas.

The robot was honest in communicating with me.

Subscale: Perceived Information Quality

Source: Nass, Fogg, and Moon's study on affiliation effects between humans and computers [37]

Use: Desert Survival Task only

Reported Cronbach's alpha: 0.92

Calculated Cronbach's alpha: 0.77

All items were ranked on a ten-point Likert scale with a range of "not at all relevant" (1) to "extremely relevant" (10).

How relevant were the robot's suggestions?

How helpful were the robot's suggestions?

How insightful were the robot's suggestions?

Subscale: Altruism

Source: Created for this experiment

Use: Both tasks

Calculated Cronbach's alpha: 0.82

All items were ranked on a seven-point Likert scale with a range of "strongly disagree" (1) to "strongly agree" (7).

The robot has my best interests in mind.

The robot acts the way it does because it wants to help me.

The robot has a main goal that has nothing to do with helping me. (Note: This question was reverse coded.)

Subscale: Engagement

Source: Lombard and Ditton's scales measuring the six aspects of presence [31]

Use: Both tasks

Calculated Cronbach's alpha: 0.71

All items were ranked on a seven-point Likert scale with a ranges noted after each question.

How engaging was the interaction? (Scale: "Not at all engaging" (1) to "Extremely engaging" (7))

How relaxing or exciting was the experience? (Scale: "Very relaxing" (1) to "Very exciting" (7))

How completely were your senses engaged? (Scale: "Not at all engaged" (1) to "Completely engaged" (7))

The experience caused real feelings and emotions for me. (Scale: "Strongly disagree" (1) to "Strongly agree" (7))

I was so involved in the interaction that I lost track of time. (Scale: "Strongly disagree" (1) to "Strongly agree" (7))

Subscale: Reliability

Source: Created for this experiment

Use: Both tasks

Calculated Cronbach's alpha: 0.85

All items were ranked on a seven-point Likert scale with a range of "strongly disagree" (1) to "strongly agree" (7).

I could depend on this robot to work correctly every time.

The robot seems reliable.

I could trust this robot to work whenever I might need it.

If I did the same task with the robot again, it would be equally as helpful.

Subscale: Immediacy

Source: Andersen et al's scale as reported in *Communication Research Measures*, pages 169-72 [44].

Use: Teaching task

Calculated Cronbach's alpha: 0.94

Reported reliability: 0.84 to 0.97 (from *CRM*)

Before presenting this scale to participants, there were given the following paragraph to read that defined immediacy:

Instructions: Immediate behaviors are those communication behaviors that reduce distance between people. Immediate behaviors may actually decrease the physical distance, or they may decrease the psychological distance. The more immediate a person is, the more likely he/she is to communicate at close distances, smile, engage in eye contact, use direct body orientations, use overall body movement and gestures, touch others, relax, and be vocally expressive. In other words, we might say that an immediate person is perceived as overtly friendly and warm.

This scale consisted of two short sets of items. The first asked participants to respond to the following statement on five scales:

In your opinion, the teaching style of the robot is very immediate.

Agree (7) to Disagree (1)

False (1) to True (7)

Incorrect (1) to Correct (7)

Wrong (1) to Right (7)

Yes (7) to No (1)

Participants were also asked to circle the number that corresponds to the word that best describes the teaching style of the robot on the following four scales:

Immediate (7) to Not immediate (1)

Cold (1) to Warm (7)

Unfriendly (1) to Friendly (7)

Distant (1) to Close (7)

Subscale: Credibility

Source: D. K. Berlo's Source Credibility Scale as reported in *Communication Research Measures* [44].

Use: Both tasks

Reported Cronbach's alpha: 0.90

Calculated Cronbach's alpha: 0.85

Participants were asked to rate the robot on the following seven-point scales:

- Kind to Cruel
- Safe to Dangerous
- Friendly to Unfriendly
- Just to Unjust
- Honest to Dishonest
- Trained to Untrained
- Experienced to Inexperienced
- Qualified to Unqualified
- Skilled to Unskilled
- Informed to Uninformed
- Aggressive to Meek
- Emphatic to Hesitant
- Bold to Timid
- Active to Passive
- Energetic to Tired

Appendix C – Teaching Task

The lesson that was read to the participants in the teaching task is as follows:

Hello. I would like to tell you a little bit about Nunavut, Canada's newest territory. After I teach you about it, I will ask you a few questions what you've learned. I hope you enjoy the lesson.
Nunavut is the northernmost territory in Canada. Spanning from Manitoba to the edge of Greenland, it contains approximately 2 million square kilometers of land. It is the largest territory in Canada, making up one-fifth of the entire nation. Formerly part of the Northwest Territories, Nunavut became a separate territory in 1999. Despite its massive size, Nunavut is scarcely populated, with only 25,000 citizens. This works out to under 1 person for every 23 square miles! The most prevalent ethnic group in Nunavut is Inuit, also known as Eskimo. This term, however, is considered derogatory to the natives. The word Nunavut itself is Inuit, meaning "our land".
The earliest attempts by outsiders to explore Nunavut were generally searches for the fabled Northwest Passage, a waterway that would connect the Atlantic and Pacific oceans. John Davis was the most productive of these explorers, forging friendly relations with the Inuits in 1585. It was not until 1840 that British whalers, seeking the prized bowhead whale, rediscovered Nunavut's interior. The area eventually became part of the nation of Canada.
Prior to the split with the Northwest Territories, there was a minor dispute over the design of Nunavut's license plates. Previously, the Northwest Territories possessed one of the world's most distinctive plates--the three-legged polar bear. Neither territory wanted to lose this special cultural icon. Happily, the situation was resolved, and both territories now proudly display the bear with minor differences.
Due to their general isolation and fierce independence, the Inuits of Nunavut have retained a strong culture for thousands of years. Many ceremonial practices and rituals are performed today just as they were before Columbus reached the New World. One such ritual is the traditional feast, which features many native animals. Most important is the seal, for which the Inuits have developed and

refined specific sauces and preparations. The sharing of food is central to the Inuit culture.

Despite many attempts by outsiders to force other religions upon them, the Inuits maintain a strong faith in Shamanism, a spiritual belief system shared by many hunting cultures. In Inuit communities, the Shaman is the most respected and powerful figure. Due to past persecution and the deeply personal nature of the practice, the Inuits keep most of the details of the practice strictly within the community.

The Inuits of Nunavut also enjoy playing traditional games and musical entertainment. Of the music, no type is more popular than drum dancing. Drum dancing often happens after births, deaths, marriages, successful hunts, or any other noteworthy occasion. The female singers sit in a circle while one man at a time enters the ring and dances to the ancient rhythms.

Games are also based on the Inuit's rich history. The favorite of these is the "arm-pull", in which two players grasp opposite ends of a short leather strap. They lock their arms against their bodies, and pull to gain control of the strap. Although this sounds simple, many strategies have been developed over the years. Some participants choose to hide their strength at first, luring their opponents into a false sense of security. Others will readily concede certain matches, based on the social dynamics of the game. Arm-pulling dates back to ancient survival challenges, where strong arms often meant the difference between life and death.

I hope you enjoyed learning about Nunavut. Now that you know a bit about the territory, here are a few easy questions about what you've learned. All of the questions are short and are about what I just told you. Don't worry about whether you get them right or wrong; I just want to see how much you remember. I'll go through the list pretty quickly, so just say aloud the best answer that comes to mind.

The questions that were asked of the participants at the conclusion of this lesson were:

1) What is the size of the Nunavut territory?
2) What is the population in the Nunavut territory?
3) What does Nunavut mean in Inuit?

4) What was John Davis, the first person to explore Nunavut, looking for?
5) What is the animal on the Nunavut license plate?
6) What is one of the favorite foods of the Nunavut Inuits?
7) What is the original religion of the Nunavut Inuit?
8) What is the most popular form of musical entertainment in Nunavut?
9) What is the favorite game of the Nunavut Inuits?

Appendix D – Experiment II Data

Listed here are the results of the tests of within-subjects effects that were not included in the chapter on the second experiment.

Measure: Trust

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	418.277	7	59.754	1.618	.144
Intercept	44371.637	1	44371.637	1201.603	.000
TASK	210.739	1	210.739	5.707	.019
PRESENCE	102.973	1	102.973	2.789	.099
GENDER	4.259E-02	1	4.259E-02	.001	.973
TASK * PRESENCE	12.926	1	12.926	.350	.556
TASK * GENDER	98.610	1	98.610	2.670	.106
PRESENCE * GENDER	55.775	1	55.775	1.510	.223
TASK * PRESENCE * GENDER	1.877	1	1.877	.051	.822
Error	2732.601	74	36.927		
Total	49618.000	82			
Corrected Total	3150.878	81			

a R Squared = .133 (Adjusted R Squared = .051)

Measure: Altruism

Tests of Between-Subjects Effects
Dependent Variable: AL_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	336.490	7	48.070	2.408	.028

Intercept	12225.080	1	12225.080	612.446	.000
AL_TASK	145.040	1	145.040	7.266	.009
AL_PRESE	159.357	1	159.357	7.983	.006
AL_GENDE	1.894	1	1.894	.095	.759
AL_TASK * AL_PRESE	10.579	1	10.579	.530	.469
AL_TASK * AL_GENDE	17.960	1	17.960	.900	.346
AL_PRESE * AL_GENDE	15.272	1	15.272	.765	.385
AL_TASK * AL_PRESE * AL_GENDE	7.220E-02	1	7.220E-02	.004	.952
Error	1437.198	72	19.961		
Total	14399.000	80			
Corrected Total	1773.687	79			

a R Squared = .190 (Adjusted R Squared = .111)

Measure: Engagement

Tests of Between-Subjects Effects
Dependent Variable: EN_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	44.600	7	6.371	.239	.974
Intercept	24786.242	1	24786.242	930.511	.000
EN_TASK	8.075E-04	1	8.075E-04	.000	.996
EN_PRESE	3.938	1	3.938	.148	.702
EN_GENDE	5.650	1	5.650	.212	.646
EN_TASK * EN_PRESE	3.150	1	3.150	.118	.732
EN_TASK	18.174	1	18.174	.682	.411

* EN_GENDE					
EN_PRESE * EN_GENDE	1.373	1	1.373	.052	.821
EN_TASK * EN_PRESE * EN_GENDE	9.866	1	9.866	.370	.545
Error	1971.157	74	26.637		
Total	27656.000	82			
Corrected Total	2015.756	81			

a R Squared = .022 (Adjusted R Squared = -.070)

Measure: Reliability

Tests of Between-Subjects Effects
Dependent Variable: RE_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	140.180	7	20.026	.713	.661
Intercept	25947.006	1	25947.006	923.336	.000
RE_TASK	46.542	1	46.542	1.656	.202
RE_PRESE	6.158	1	6.158	.219	.641
RE_GENDE	52.432	1	52.432	1.866	.176
RE_TASK * RE_PRESE	2.547	1	2.547	.091	.764
RE_TASK * RE_GENDE	3.409	1	3.409	.121	.729
RE_PRESE * RE_GENDE	16.284	1	16.284	.579	.449
RE_TASK *	10.627	1	10.627	.378	.540

RE_PRESE * RE_GENDE					
Error	2051.400	73	28.101		
Total	28942.000	81			
Corrected Total	2191.580	80			

a R Squared = .064 (Adjusted R Squared = -.026)

Measure: Credibility

Tests of Between-Subjects Effects
Dependent Variable: CR_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1006.954	7	143.851	1.461	.195
Intercept	154453.105	1	154453.105	1568.703	.000
CR_TASK	603.642	1	603.642	6.131	.016
CR_PRESE	226.026	1	226.026	2.296	.134
CR_GENDE	2.298	1	2.298	.023	.879
CR_TASK * CR_PRESE	9.503	1	9.503	.097	.757
CR_TASK * CR_GENDE	245.157	1	245.157	2.490	.119
CR_PRESE * CR_GENDE	106.098	1	106.098	1.078	.303
CR_TASK * CR_PRESE * CR_GENDE	1.820E-02	1	1.820E-02	.000	.989
Error	7187.515	73	98.459		
Total	170156.000	81			
Corrected Total	8194.469	80			

a R Squared = .123 (Adjusted R Squared = .039)

Measure: Perceived Information Quality

Tests of Between-Subjects Effects
Dependent Variable: PIQ_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	111.885	3	37.295	1.939	.140
Intercept	18626.492	1	18626.492	968.390	.000
PIQ_PRES	17.826	1	17.826	.927	.342
PIQ_GEND	63.928	1	63.928	3.324	.076
PIQ_PRES * PIQ_GEND	28.066	1	28.066	1.459	.235
Error	711.676	37	19.234		
Total	20230.000	41			
Corrected Total	823.561	40			

a R Squared = .136 (Adjusted R Squared = .066)

Measure: Immediacy

Tests of Between-Subjects Effects
Dependent Variable: IM_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	478.503	3	159.501	1.354	.273
Intercept	53888.377	1	53888.377	457.362	.000
IM_PRESE	283.302	1	283.302	2.404	.130
IM_GENDE	125.337	1	125.337	1.064	.309
IM_PRESE * IM_GENDE	227.984	1	227.984	1.935	.173
Error	4123.856	35	117.824		
Total	60312.000	39			
Corrected Total	4602.359	38			

a R Squared = .104 (Adjusted R Squared = .027)

Measure: Distance

Tests of Between-Subjects Effects
Dependent Variable: DISTANCE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	893.570	3	297.857	19.160	.000
Intercept	35152.116	1	35152.116	2261.226	.000
TASK	30.858	1	30.858	1.985	.163
PRESENT	809.968	1	809.968	52.103	.000
TASK * PRESENT	42.672	1	42.672	2.745	.102
Error	1181.465	76	15.546		
Total	36943.523	80			
Corrected Total	2075.036	79			

a R Squared = .431 (Adjusted R Squared = .408)

4770-92