ABSTRACT

BIWEN, ZHU. Design of Etiquette for Patient Robot Interaction in a Medicine Delivery Task. (Under the direction of Dr. David B. Kaber).

Robots are currently being developed to perform medicine delivery tasks for patients in healthcare environments (e.g., hospitals and nursing homes). It is important to ensure that robots are able to interact with patients in ways that patients find comfortable and that promote perceptions of quality health care. Two interface features in robot design relevant to patient-robot interaction (PRI), including robot language strategies (derived from Brown and Levinson's etiquette model in human-human interaction) and robot physical appearance, were assessed in this study for affects on perceived robot social etiquette.

The goal of the study was to determine if Brown and Levinson's model of linguistic etiquette could be extended to the human robot interaction (HRI) domain, and how different etiquette strategies might affect performance of humans and robots as mediated by manipulations of robot physical features, in a simulated concurrent medicine delivery task. To achieve this goal, a humanoid robot prototype and a neutral-looking robot prototype were used in the experiment to deliver medicine reminding utterances to subjects following different etiquette strategies. This occurred while subjects performed a primary cognitive task (i.e., the robot interaction was considered a disruption).

Results showed that the etiquette model could only be partially extended to the HRI domain. Subjects interpreted and understand a robot negative utterance strategy consistent with the original etiquette model. However, such consistency was not found for a utterance following a positive etiquette strategy. Results showed that a negative etiquette strategy promoted both user task performance and robot performance (in terms of user compliance to

robot requests), and resulted in the highest user perceptions of perceived etiquette (PE). A positive etiquette strategy led to the lowest PE ratings and should be avoided in social service robot interface design. It was also found that robot utterances with higher etiquette levels did not necessarily lead to higher subjective PEs.

With respect to the effect of robot appearance on PE ratings and performance, it was found that user task performance scores were higher under the humanoid robot condition as compared with a neutral-looking robot, but no differences were found in terms of PE. Users also responded faster to robot requests in the humanoid robot condition. These results provide a basis for determining appropriate etiquette strategy types and robot appearance to promote better collaborative task performances for future healthcare delivery applications of service robots.

Design of Etiquette for Patient Robot Interaction in a Medicine Delivery Task

by
Biwen Zhu

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Industrial and Systems Engineering

Raleigh, North Carolina

2009

APPROVED BY:

____
David B. Kaber
Committee Chair

____
Jeff Thompson
Member

____
Haig Khachatoorian
Member

BIOGRAPHY

Biwen Zhu grew up in Shanghai, China, and attended Shanghai JiaoTong University, where he received dual Bachelor degrees in Industrial Design and Animation. He entered the Ergonomics lab in the Department of Industrial and Systems Engineering at North Carolina State University in fall 2007. His research interests include human robot interaction, human computer interaction, web 2.0 and user experiences. He plans to continue his PhD research after completion of his Master’s degree.

.

ACKNOWLEDGMENTS

I would like first to thank Dr. Kaber for his guidance and continuous support for my two-year study at NC State. It was a great learning experience working with him in the past few years. In addition, I would like to thank Dr. Couch, Dr. Khachatoorian, and Dr. Thompson for their valuable inputs and suggestions on my thesis project. Finally, I greatly appreciate my parents for their unconditional love and encouragement.

I would also like to thank The Ergonomics Center of North Carolina (ECNC) for supporting my Master's study over the past two years. This research would not have been possible without the Center's funding to the Ergonomics Lab.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

In recent years, a shortage of registered nurses (RNs) has been reported in the U.S. (Janiszewski Goodin 2003). This shortage is expected to extend to 2020 with an estimated 400,000 RN vacancies (Murray 2002). With an aging population and ever increasing number of patients in hospitals, there is a high demand for patient services, such as medicine delivery. Nurses typically perform medicine delivery on a daily basis and this contributes to a high workload, which may lead to errors and slips. The probability of making an error in medicine delivery tasks has been observed to dramatically increases when work shifts for nurses are longer than 12 hours (Rogers, Hwang et al. 2004). Such errors can include administering an incorrect amount of medicine to a patient, misreading a prescription, or giving a patient the incorrect medicine. For example, about 1,200 people died in public hospitals in Britain in the year 2000 because of mistakes in prescribing and administering medicine (Lyall 2001). A more recent accident happened in 2008 (Parker 2008). Infant twins were given an adult dose of a blood thinner and almost died because a nurse had misread a label on a medicine bottle. From these stories, it can be seen that nursing errors can lead to serious injuries or even fatalities.

Nurses are the largest group of health care professionals that provide direct patient care in hospitals. Therefore, nursing performance is tightly linked to the overall quality of health care perceived by patients (Hassmiller and Cozine 2006). Due to the nurse shortage and the importance of nursing, there is a need for development of programs for recruitment and retainment of RNs. Such programs may also serve to improve the overall image of the nursing industry (Janiszewski Goodin 2003). Legislation has also been proposed to help rectify the nursing shortage. These suggested solutions, however, require great amounts of time to develop and they may not be that effective in the short term.

As an additional potential solution to the nursing shortage, improvements in intelligent control systems and precision sensors have fostered a variety of robot applications in the healthcare field. Such technology may provide another way to deal with nurse workload and potential human errors. Robots can be used to assist nurses with routine tasks. Example nursing robots include the Helpmate, which is used for medicine delivery tasks (Krishnamurthy and Evans 1992), the humanoid robot Pearl for cognitive reminding tasks (Pollack, Engberg et al. 2002), and the assistive robot DO-U-MI for walking tasks (Park, Hong et al. 2001), etc. An overview of socially interactive robots in other domains can also be found in Fong et al. (2003) and Zhang et al. (2008).

Robots used for delivery tasks, in particular, can automatically perform point-to-point navigation within hospitals, carrying medicines, meals, medical records or lab specimens to nursing stations. Some of them follow preprogrammed routes with the capability of taking elevators and opening electronic doors; whereas, few others autonomously navigate in the environment using natural landmarks. In the near future, it is highly possible that these robots would be enhanced to perform delivery tasks directly to patients. For such applications, it is important to ensure that robots are able to interact with patients in ways that patients find comfortable in order to promote perceptions of healthcare quality.

Faced with a stream of sensory data, a socially competent robot should know what data matters most to the current task, what action should be taken first, and what to learn and how to learn in human environments (Breazeal 2004). Furthermore, a robot's ability to be polite and show respect to patients is also considered important to success in patient service, which is especially true in the initial cycle of patient-robot interaction (PRI).

This concept is consistent with what Miller (2002) defined as "etiquette" in the context of human-computer interaction (HCI). He said computers should exhibit what are considered to be socially polite behaviors in specific contexts to support collaborative task performance. Several studies have been conducted relating to the design and evaluation of social etiquette agents or systems, particularly for pedagogical systems (Tzeng 2004; Wang, Johnson et al. 2005; Wilkie, Jack et al. 2005; Tzeng 2006). However, much less research has been done to study "etiquette effects" in human interaction with social service robots. In addition, although the field of HCI might offer some insights for designing optimum interaction experiences for human robot interaction (HRI), significant differences exist between the two domains (Breazeal 2004). A robot has a physical embodiment that can provide additional affordances during the interaction. Therefore, there is a need to conduct studies to investigate differences in user experiences when interacting with socially polite robots compared with socially impolite robots. Etiquette is a fairly broad concept in agent-agent interaction (which will be defined later). The present study is limited to one dimension of etiquette, specifically robot linguistic etiquette.

One aspect of robot interface design important to HRI is how a service robot should physically look relative to human appearance. There is no general agreement on this in the literature. Severinson-Eklundh et al. (2003) argued that in home environments, service robots with a tall and humanoid shape would make users more uncomfortable than ones with a short, neutral cylindrical shape. Robots are machines and extra humanoid features may generate false expectations of behavior or even fear. However, some other researchers argued that a humanoid form of service robot would ease interaction because rules for human social interaction might be evoked, and thus, humanoid robots might support more intuitive interaction (Breazeal and Scassellati 1999). Furthermore, development of appropriate mental models of robot capabilities might be easier if robots have certain human characteristics. Hinds et al. (2004) suggested that humanoid robot may be appropriate when tasks are complex or risky, because people will readily delegate responsibility to them. Machine-like features are appropriate when the robots are expected to be unreliable, less well-maintained for tasks or in situations where personal responsibility should be emphasized. In support of humanoid robots, Kanda et al. (2004) conducted a field study to examine the extent to which a humanoid robot could develop a relationship with children in a language-learning scenario. They found that humanoid features in a robot helped establish common social goals and sustain long term social relationships.

In human machine interaction, Nass and Moon (2000) found that humans often respond to computers and other technologies using the same social rules as they use in human-human interaction (HHI), and the extent to which social rules are applied depends on the number and strength of human-like cues conveyed in the technology. In particular, Nass et al. (1999) found that human evaluation of a computer agent's performance was more positive if the human had interacted with the same agent before. This phenomenon implies that humans may even show politeness towards computers.

In the current study, the influence of robot appearance, in addition to language etiquette, on user perception and performance of concurrent tasks was examined. The focus of this investigation centered on physical and behavioral (language) aspects of robot design, as compared to etiquette in computer interface design.

In order to provide a basis for conducting this study, a detailed discussion of relevant literature is provided in the following sections. The review of literature is organized into the following major parts: formal definition of etiquette and review of its origin in HCI studies (Section 2.1); and description of a traditional model of linguistic etiquette in HHI (Section 2.2). Detail on the assumptions and the properties of the model, as well as its applications in HCI domain are also provided in this section. Following this, several studies regarding the design and evaluation of etiquette in HRI are reviewed and summarized in Section 2.3. Emphasis is placed on the differences that exist in etiquette studies in HRI compared with those in HCI. Finally, in Section 2.4, studies on anthropomorphism and humanoid form in robot design are reviewed and summarized. A definition of anthropomorphism is provided and its influences on user perception of robot performance (particularly robot etiquette) are also discussed.

2. LITERATURE REVIEW

2.1. Etiquette and linguistic etiquette

Etiquette, as defined through Wikipedia, is “a code that influences expectations and social behaviors, according to contemporary conventional norms within a society, social class, or group”. Rules of etiquette are usually learned through experience in a community (Preece 2004) and are culture-specific. For example, in China, one is not expected to take the last item of food from a common plate or bowl without first offering it to others at the table. Such behavior is regarded as impolite or even insulting to the generosity of a host. In the US, however, the guest is expected to finish all the food provided, as an indication of how much he/she likes the food. Etiquette, according to this definition, refers to the *expected moves* in context that allow participants to make inferences during the interaction, and is different (broader) from politeness (Miller, Wu et al. 2004). Politeness is a linguistic term which usually deals with the communication process. Miller et al. (2004) said politeness is more like a *method* by which we signal, interpret, maintain and alter power relationships, familiarity relationships and interpretations of the degree of imposition of an act.

Linguistic etiquette, a focus of this project, is the set of expectations of language usage in a social group. In the review below, justification is provided for the study of etiquette in advanced human machine interaction. The origin and a formal definition of etiquette, as well as its characteristics and influential factors, are also provided. Finally, a tentative list of etiquette rules in nursing (specific for medicine delivery tasks) is summarized at the end of the section.

2.1.1. Origin of etiquette in HCI

The motivation for introducing and developing the concept of HCI etiquette arose during research on a project called the Rotocraft Pilot's Associate (RPA), which was to design and implement a computer associate to assist helicopter crews in identifying relevant operational information during flight (Miller and Funk 2001). Previous automated solutions (e.g., the Cockpit Information Manager (CIM)) were concerned with identifying ways to present flight information that would not require active input from the user. The CIM system, in particular, recognizes crew actions and senses related context information, infers a goal and allocates a task for the pilot. The appropriate information is then sent to a cockpit display. Pilots can also control the set of options that the CIM is permitted to consider, and apply preference weights to those options, during initial configuration of the system. The advantage of such a configuration is a decrease in the human operator's cognitive workload during flight. However, Miller and his colleagues argued that this system failed to behave well in accordance to the so called "*flight operational etiquette*". For example, it could not report on its activities, its perception of the activities of others, or take instruction about the activities it should be engaged in. Users were often frustrated if the system misunderstood their needs and presented the wrong information. Such frustration became even worse when the pilot could not redirect the action of the computer. The authors also found that pilot expectation of the behaviors for the computer associate in this context included answering questions about what the agent's current goals were, providing easy ways to override and correct the associate errors, and trying not to interrupt the user.

In order to achieve these etiquette expectancies, a four-button "Crew coordination and task awareness" display was prototyped as a simple "meta-communication" control mechanism between the pilots and the associate, which was further integrated in the main CIM system. The buttons were used to report the associate's varied inferences on the cockpit state in textual form (indicating what the associate thought was going on in the cockpit). By pressing the buttons, the interface permitted the pilot to override the CIM's current inferred tasks and insert new ones. The authors posited that the automation of the flight deck demanded that the associate at least provide this level of meta-communication ("etiquette") about its intents and its knowledge of the intent of others to efficiently and effectively support performance.

Extensive full mission simulations were conducted to evaluate the updated CIM behaviors. Each crew flew 14 missions in the experiment, seven with the full RPA aiding (CDAS), and seven with an Advanced Mission Equipment Package (AMEP), only. Results showed that in general, pilots found the CIM behaviors to be "of Use" or "Of considerable use". In evaluating the pilots-reported frequency of overrides and self-corrections of CIM behaviors, the average fell between "frequently" and "now and then". Workload scores were significantly lower for the CDAS condition than AMEP condition. Pilots found that they were more effective with CDAS than without. The authors wondered how such perceived effectiveness and usefulness of the CDAS could be achieved along with perceived error rates. They posited that this may have been due to the "etiquette" embedded in the updated CIM system. This outcome supported the hypothesis that the capability for the crew to interact directly with the associate regarding the active task was a capability that the pilots welcomed, and that would enhance the overall effectiveness and usefulness of the CIM.

To further understand the need to consider etiquette in human machine interaction, another study was performed by Parasuraman and Miller (2004). This research was conducted to determine whether the norms of human-human etiquette affect the calibration of human trust in the usage of highly-critical automation systems. A flight simulator was used in the experiment as a platform for presenting automation. An experiment was conducted, in which the authors manipulated two independent variables: human-machine etiquette and system reliability. Target responses were user overall performance and rated trust in automation. The authors defined good automation etiquette as a communication style that was “non-interruptive” and “patient”. Etiquette automation support was provided after a 5-second warning and not at all if an operator was already performing the requested action. Bad automation etiquette was interruptive and impatient. It provided advice without warning and urged users to the next query before they finished with the first one. Two levels of reliability were defined: high (8 correct pieces of advice out of 10) and low (6 correct pieces of advice out of 10).

Results showed that the majority of subjects were aware of the etiquette automation manipulated in the trials. Good etiquette automation significantly enhanced diagnosis performance, regardless of the system reliability level. Furthermore, good etiquette compensated for performance degradations caused by low automation reliability. When evaluating the response of user trust in automation, the authors found that high reliability increased trust ratings significantly, so did good automation etiquette. The authors suggested that etiquette should be considered in human machine interaction, especially in critical and highly technical automation domains.

2.1.2. Definition of etiquette in HCI

Based on the experiment outcomes of the RPA project, Miller developed the concept of *etiquette* in HCI (Miller 2002) and identified it as an important factor facilitating smooth and effective interaction between humans and computers. Etiquette, as the author defined, is “the set of prescribed and proscribed behaviors that permit meaning and intent to be ascribed to actions in a common social setting” (Figure 2-1). Etiquette can be general or domain specific (specialized etiquette). The former is often referred to as general politeness or social niceties, which are more likely to be used in situations where a specific context or relationship is unknown (e.g., interacting with strangers). On the other hand, when we are talking about computers, which are always used as a tool to assist people in achieving certain goals in specific work domains, specialized codes of etiquette should be adopted for the computer to use. In addition, such domain and context-dependent etiquette may involve substantial deviation from general “polite” forms. Different etiquette sets could also have intersections with each other (Figure 2-2). These etiquette sets can be used for facilitating group identification, fostering communication and generating expectations.

Related to Miller’s work, there are also other relevant terms defined in the literature. For example, Hayes et al.(2002) referred to “Human Computer Etiquette” as a type of mental “affordance”, which defines the rules of behaviors necessary for enabling computers to act as socially acceptable members of a larger team working together towards a common goal.

Whitworth (2005) provided a full definition of politeness in interaction, specifically “any unrequired support for situating the locus of choice control in social interaction with another party to that party, given that the control is desired, rightful and optional”. He argued that giving/taking choice is central to the distinction of polite/impolite behaviors, and such behavior is considerate and polite only in the situation that a human wants the choice. This concept is different from the concept of etiquette for the following reasons. Firstly, politeness is a common theme of social niceites and is different from etiquette, which is more focused on specific behaviors and varies between cultures and contexts. Second, etiquette is a set of expected behaviors in a common social setting; whereas, politeness should be considered as a method to achieve such expectancies. In this study, Miller’s definition of etiquette was considered and extended to advanced human robot interaction, as this definition has been most widely used in the past literature.

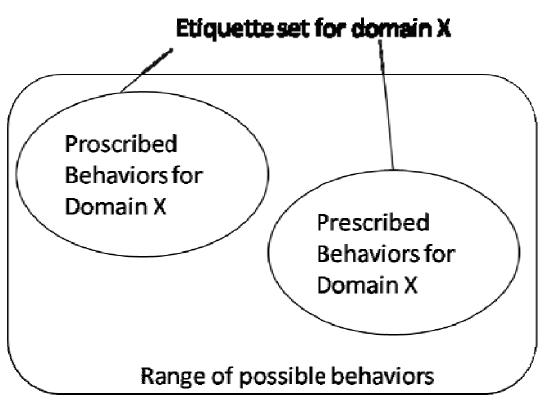

Figure 2-1. Etiquette definition for domain X (reproduced from Miller, 2002)

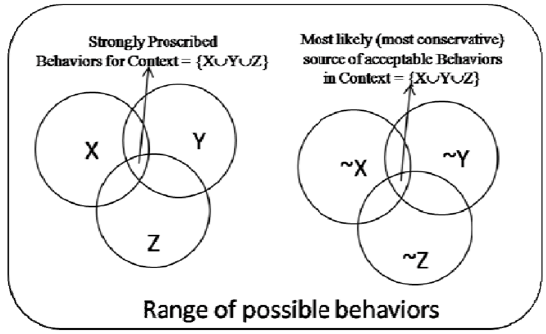

Figure 2-2. Different etiquette for different interactions (reproduced from Miller, 2002)

2.1.3. Characteristics and influential factors in etiquette

Characteristics and influential factors for HCI etiquette have been identified and discussed in detail by Miller (2002), five of which are relevant to the current study. These are illustrated in Figure 2-3. First, etiquette is context and role based. Different contexts and different types of interactions have different etiquette associated with them. The appropriate etiquette is a function of etiquettes for the set of overlapping contributing contexts. Within the same domain/context, the acceptable behaviors for one role may be strongly forbidden for another role (e.g., pilot and copilot). Power relationships are frequently encoded as roles. Behaviors that violate role boundaries are strongly marked and potentially confusing. Second, etiquette relaxes and evolves as we become more familiar with the person and the context. Relaxed etiquette might be more efficient during the interaction. Third, etiquette is only expected of intelligent agents (e.g., robots). As a system becomes more complex, adaptive, and autonomous, the need for the system to exhibit proper etiquette increases. A human's expectations for such agents also increases. Fourth, etiquette in human-human interaction is functional and is largely homogenous across societies and cultures, but its forms are arbitrary and vary across cultures. For example, people with different cultural backgrounds have different ways of exhibiting good table manners; however, their functional goal of showing respect to guests at the same table is the same. Lastly, etiquette constraints are "soft" constraints. Violations of etiquette rules may not always result in a failure of a goal, task or design. On the other hand, adherence to the rules makes interaction smoother, more pleasant and efficient. Each of these characteristics is addressed in the design of experiment for this study (see Section 4).

An advanced robotic system was chosen as an intelligent agent in order to study etiquette in HRI. The robot was designed for use in the context of health care environments and conducted medicine delivery tasks for patients in the experiment. The robot's role was defined as a nursing staff member. An interaction scenario was developed and no prior experience with robots was set as a criterion for subject recruitment. The relevance of etiquette characteristics to the present study is summarized in the figure below.

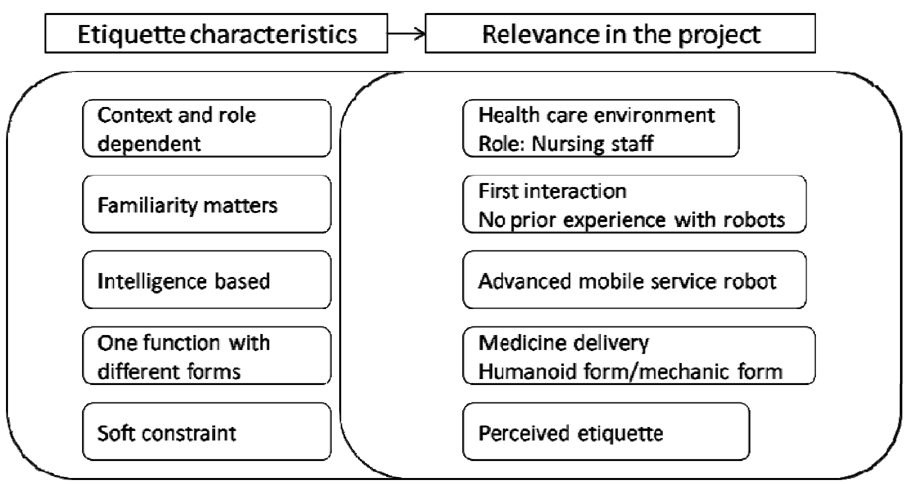

Figure 2-3. Etiquette characteristics and relevance (adopted from Miller, 2002)

Based on these characteristics, it can be seen that etiquette design in HCI is quite different from traditional usability design, which only focuses on a single individual and his/her computer. The etiquette concept applied to a human-computer relationship emphasizes the role that an individual (and computer) plays in a greater team, social and organizational context. The goal of etiquette is to foster rich communication and to build more social capital in an existing system (Miller 2002). Hayes et al. (2002) also argued that HCI etiquette should be a special category of general rules for effective HCI design. Several reasons include: (1) the rules for effective HCI design focus on designing interfaces that respect the limitations of human perception, memory, attention, or reaction speed, while HCI etiquette rules focus on satisfying social interaction constraints and conventions developed within a national culture or specific work group; and (2) user and machine roles have a major influence on etiquette because they determine power precedence in work groups, while traditional HCI design focuses more on the tasks, not roles. A simple pictorial representation for the difference between HCI usability and etiquette is provided in Figure 2-4.

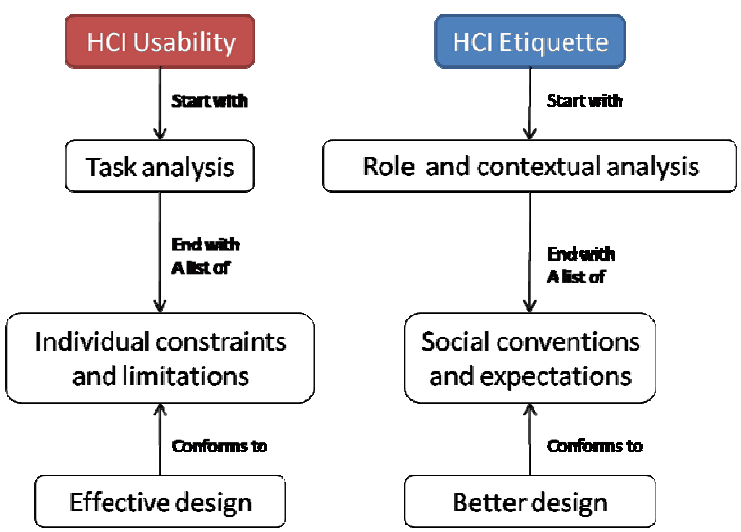

Figure 2-4. Comparison of HCI Usability and HCI Etiquette

2.1.4. Why consider linguistic etiquette in human-agent interaction?

There are several major reasons why linguistic etiquette is emphasized in this study. First, as mentioned earlier, etiquette expectations form a user's perception of the sociability of machines in context. Different sets of etiquette rules need to be studied separately in order to understand how they influence user perception of robot sociability in context.

Second, Nass (2004) proposed a list of factors that may trigger etiquette responses in HCI. He said that agents or entities having certain features are more likely to elicit both broader and more robust polite responses and polite expectations from users. Among these factors, language use comes first. Language or speech has been evaluated as a feature for causing user perception of humanness in interacting with machines. The "*Turing Test*" is a test of a computer's artificial intelligence and articulation of output for users. In this test, one substitutes a computer for one of two human participants in a keyboard and screen dialogue. The objective is to determine whether the computer substitution is detected by the remaining human participant. It was argued that proper language use is an influential factor in discriminating a human from a machine in this test. In addition, some other factors, like emotion manifestation and engagement with the user, are related to language use. Potential causal relationships exist among the factors.

Thirdly, HRI experiments have been conducted including humans interacting with a medicine delivery robot (Zhang,et al., 2008). In this research, participants were asked some etiquette related questions to get feedback on various robot configurations. One question was as follows:

“We are going to integrate some new features/functions in a medicine delivery robot to make it more “polite” towards users. What behaviors would you expect from the robot to ensure good etiquette?”

It was found that the majority of the participants thought voice, speech and communication were key features in assessing robot etiquette. They expected a robot to be sociable in terms of speech. For example, the robot should introduce itself first when it makes entry into a patient room, engage in questions and answers with the patient and be cheerful in tone. Clear and precise communication was also expected by a minority of participants (2 of 24).

Finally, there exists a linguistic model of human-human etiquette in the social-linguistic area. Brown and Levinson’s “face-saving” model of politeness (Brown and Levinson 1987) has been adopted and used extensively in etiquette studies in HCI (to be reviewed later). Extension of the model to the realm of human (or patient) robot interaction is, however, a new research area and might have many implications for ensuring optimum user experiences during interaction with social service robots.

For these reasons, linguistic etiquette was selected as the focus in this project. The relationship between etiquette and linguistic etiquette is illustrated in Figure 2-5. As mentioned earlier, etiquette itself is a set of expected behaviors, and politeness is the method used to help achieve etiquette. Thus, we can say, in a more specific way, that Brown and Levinson’s model of politeness may be used as a method to achieve good linguistic etiquette during human-robot interaction.

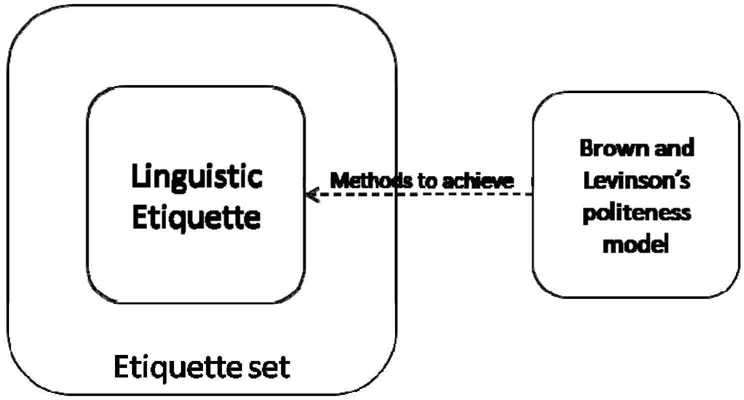

Figure 2-5. Etiquette set and linguistic etiquette set

2.1.5. Etiquette in nursing

Before concluding this section and proceeding with the description of the etiquette model, a general coverage of etiquette expectations for nursing staff, especially in medicine delivery tasks, is required to better understand the concept of “etiquette” in real environments. The nursing profession has its distinct code of ethics and professional etiquette. In this domain, factors such as neatness, punctuality, economy, and quietness fall under etiquette (Perry 1906). For example, a nurse’s neatness ensures her surgical cleanliness, which is a fundamental principle of all operations. The time and manner of administration of drugs must be accurate and recorded clearly. Economy in task performance includes thoughtful consideration of use of resource and avoiding wastefulness, especially towards patients. A quiet manner is also important to allow sick persons to rest.

Johansson, Oleni et al. (2002) found that patients expect nurses to play a central role in their care, act as a companion and advisor, and posses an empathetic ability and good communication skills (here, “good” means clear and straightforward). They also found that interpersonal relationships with nurses are highly valued and expected among patients. A good relationship between patients and care-givers, the characteristics of which might be mutual understanding, respect, trust, honesty or humor, is important for patient satisfaction in health care operations. For example, if a nurse takes the time to listen to a patient, as well as demonstrates an interest and commitment to contact with the patient, this may contribute to patient satisfaction. Besides this, it was also found that patients expected active nurse participation and involvement in the caring process. All of these findings represent general patient etiquette expectations in nursing.

Kahn (2008), in his article on "etiquette-based medicine", developed a tentative checklist of behaviors to ensure good patient-doctor/nurse etiquette relationships based on an informal meeting with a hospitalized patient. He posited that patients would care less about whether their doctors/nurses were reflective and empathic than whether they were respectful and attentive. Kahn's list includes: (1) asking permission to enter a patient room; (2) waiting for an answer; (3) introducing oneself, showing ID; (4) shaking hands (wearing gloves if needed); (5) sitting down and smiling appropriately; (5) briefly explaining one's role in the health care process; and (6) asking a patient how he or she feels. This list could be used in a variety of task scenarios, including medicine delivery. However, in the nursing context, this concept of etiquette has been applied in tasks that are more related to general nursing ethics and regulations.

Such expectations for human nurses may not always be important when extended to nursing robots. For example, how could one define cleanness for a robot nurse? Does anyone expect a robot nurse to wear a nursing uniform? Formal experimentation is needed to test how the patterns of human expectations for nursing robots change as compared to those for human nurses. For example, human expectations of how delivery robots should ask for a patient's attention during a service task were a focus of this study.

In summary, etiquette is a set of rules and expected behaviors in a specific context and task domain. Etiquette design in HCI has the potential to enhance system performance, for example, in the aviation domain. Etiquette is different from politeness, which is regarded as a common method humans use to achieve certain etiquette expectations in language communication. As a subset of general etiquette, linguistic etiquette is important to intelligent agent interaction and needs further investigation in the context of PRI. In medicine delivery tasks, the proper use of robot language/speech is expected to be important to interaction, based on the results of earlier study. Hence, finding proper etiquette strategies that humans commonly use in communications may be helpful for the PRI domain.

2.2.A model of human-human etiquette for linguistic politeness

A major class of social-linguistics studies has been focused on the modeling of "politeness" behaviors in human-human interaction. One of the most influential models is Brown and Levinson's "face-saving" theory of politeness (Brown and Levinson 1987). Brown and Levinson attempted to formulate a theory of how individuals from different cultures produce linguistic politeness. They assumed that every individual has two types of faces, positive and negative. A positive face is the human's desire for his value/wants to be appreciated and approved in a social setting. The negative face is the human's desire for freedom of action and freedom from imposition. It was argued that in order to achieve common communicative goals, humans assess and minimize the dangers of face-threatening acts (FTAs) during interaction by choosing appropriate linguistic strategies. For example, Brown and Levinson proposed five levels of politeness strategies (Figure 2-6). Factors influencing the choice of these strategies were also examined. (A detailed discussion is provided later in this section.)

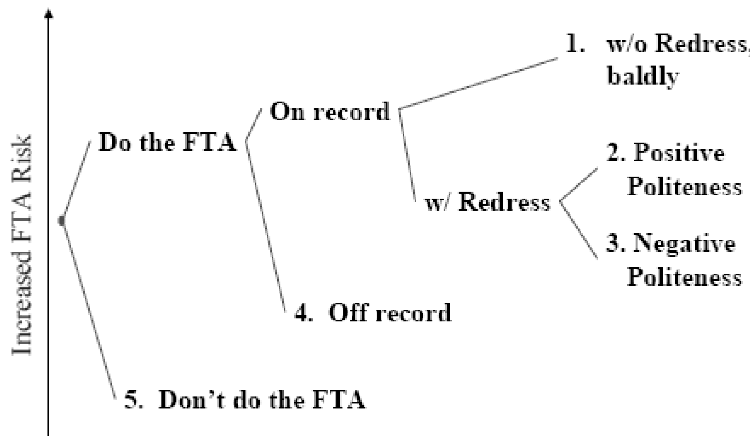

Figure 2-6. Etiquette model from Brown and Levinson (1987)

2.2.1. Assumptions of the etiquette model and properties of interactants

In Brown and Levinson's model, types of "faces" and rationality were assumed as two properties that all competent adult members of a society have. First, the notion of "face" was derived from the work of Goffman (1967) and further defined as, "the public-image that every member wants to claim for himself". The two types of face embedded in this concept were negative face and positive face. Negative face was the basic claim to territory, personal preservation, and right to non-distraction; whereas positive face concerns a consistent self-image or personality. Face is considered to be something that is emotionally invested, and can be lost, maintained or enhanced and must be attended to constantly during interaction. The authors posited that mutual knowledge of a members' social image or face, as well as the necessity to orient oneself to it in interaction, is universal across all cultures, although the content might be different. Second, the rationality assumption states that every social human being has certain rational capabilities that serve to guarantee goal attainment or “ends” in interaction and dictate the means used to satisfy the ends in interaction. It is also assumed that a rational person has the ability to weigh different means to an end, and choose the one that most satisfies the desired goals as well as save face. In this theory, the main desire of a human is to not waste effort in the interaction towards satisfying goals.

2.2.2. Definition of face-threatening acts

Given the assumptions of the universality of face and rationality, it is possible that virtually all interactions between social agents involve some degree of face-threatening acts (FTAs), acts that by their nature run contrary to the face needs of the observer or the speaker (i.e., they would lead to a loss of face). Such acts can occur via verbal or non-verbal communication, which might include speech, movement, gestures, facial expressions, etc. For example, some communicative acts, such as requests and offers, can threaten the hearer’s negative face, positive face, or both (Wang, Johnson et al. 2005). Consider a teacher’s critique of student work such as, *“Your work does not meet my requirements. Redo it now.”* Here, there is no attempt made to use politeness to mitigate a face threat. Both types of face threat exist in this example: the criticism of the student’s work is a threat to positive face, and the instruction of what to do is a threat to negative face. What is important to note here is that the relationship between FTAs and (personal) politeness is often different from the relation of FTAs to task performance. In fact, in many situations FTAs may lead to increased responsiveness of an observer and in some cases, observers may prefer communications that would constitute FTAs in Brown and Levinson's model. This observation is relevant to the experiment results reported later.

2.2.3. Strategies for mitigating the threats of FTAs

Five levels of politeness strategies in language use were proposed by Brown and Levinson in order to mitigate the effects of FTAs including: (1) simply do not do a face-threatening act (FTA), if the estimated risk of face-loss to the addressee is great; (2) using an off-record strategy, i.e., doing the act with a sort of "plausible deniability" by means of innuendo and hints. For example, if you want to ask for salt from someone at the same table, you might say *"I find the soup a little bland"*; (3) using strategies aimed at supporting and enhancing the addressee's positive face. Exaggerating interest and approval are good methods to achieve this goal. In the previous table manner example, you can say *"The soup tastes great, but I think I need more salt in it"*; (4) using strategies to address the hearer's negative face. Offering apologies and deference are common ways to achieve this. For example, you can say *"Excuse me, but could you please pass the salt for me"*; and (5) do the FTA baldly with no redressive actions, if the estimated risk of face-loss to the addressee is minimal (e.g., *"pass me the salt").* This strategy normally occurs in conversations between old friends, where etiquette expectations are relaxed and situations are more urgent. One point that needs to be clarified here is that the "bald" strategy does not mean "bold" or strong linguistic behavior. It refers to the use of a straightforward utterance strategy void of other features, such as those included in the positive and negative strategies. Brown and Levinson also examined 15 sub-strategies for positive politeness (such as claiming of common ground, seeking agreement, intensifying interests in the observer, etc.), 10 for negative politeness (e.g., being pessimistic, apologizing, minimizing imposition, etc.) and 15 for off-record strategy (e.g., giving hints, using metaphors, being ambiguous, etc.). A large set of examples were provided for each sub-strategy for three different language/culture groups.

For the current project, instead of examining all five linguistic strategies, only the last three were investigated. The first two strategies were not considered for the following reasons: (1) the "off record" strategy is highly indirect and context dependent, which is difficult or even impossible for machines to produce (Miller, Wu et al. 2004); and (2) the strategy of simply do not do the FTA would result in an incomplete robot task. This would not support the planned experiment to assess patient-robot interaction in medicine delivery tasks, as part of the project.

2.2.4. Factors influencing the choice of politeness strategies

In order to decide which level of politeness strategy should be used during an interaction, one needs to first estimate the severity of a FTA. Brown and Levinson posited that three socio-cultural variables should be considered in this process: (1) the power that the observer has over on the speaker – P (O, S); (2) the social distance between the two interactants – D(S, O); and (3) the ranked imposition of the act itself – R(x). The authors developed a descriptive model, as follows:

$$W(x) = D(S, O) + P(O, S) + R(x)$$

In this model, W(x) is the weight or severity of the FTA. The greater the value W(x) is, the more likely an interactant is to pick strategy Type 1 (simply do not do the FTA), since it would pose greater danger for the interaction. The lower the value of W(x), the more likely an interactant is to pick strategy Type 5 (bald with no redressive actions). Social distance is roughly the reverse of familiarity, or might be based on other factors such as membership in the same family organization, or social groups. This dimension is symmetrical, which means the observer and the speaker share the same social distance. The power dimension is based on the context of the interaction and differs among cultures and organizations. It is an asymmetrical metric; for example, a graduate advisor maintains certain power over his students and such power is not reversible between the two interactants. Such power also depends on the context. For example, after graduation, this "power" does not exist. The ranked imposition of the raw act is culturally and situationally defined and would depend upon the roles and duties of the parties involved. The act of a tutor correcting a student is less of a face threat/imposition for the student, since such correction is expected from a tutor. The formula presented by Brown and Levinson is not intended to give a quantitative measure of FTA severity, according to D, P and R, but is used to indicate the reasons for choosing one strategy over another.

2.2.5. Applications of the etiquette model in HCI research

Brown and Levinson's model of linguistic etiquette has been adopted in designing and evaluating pedagogical software agents. For example, Wang, Johnson et al. (2005) implemented two of the politeness strategies, direct (Type 5) and indirect (Type 2), in a computer-based learning system. They wanted to assess the effects of these two particular strategies on student cognitive and motivational factors in learning. In general, the authors found that the polite tactic influenced student perceptions of difficulty in learning the material and made tutors offering of help seem less intrusive. In the polite group, student self-efficacy improved significantly compared to students from the direct group. Students in the polite treatment scored a little higher than students in the direct group with respect to learning outcomes. Results also showed that highly extrovert students tended to agree more that the tutor was friendly and that their performance improved, in the polite treatment.

Based on the above findings, Mayer, Johnson, et al. (2006) conducted another study aimed at developing research-based principles for the design of computer-based intelligent tutors that are socially sensitive and motivating for learners. The authors adapted Brown and Levinson's politeness model and proposed eight types of etiquette strategies/statements (e.g., "direct commands"). Forty-seven college students with varied computer experience were recruited for the experiment. Overall, students rated "direct commands" and "commands attributed to computers" as lowest in negative and positive politeness; whereas, "guarded suggestions" were rated as highest. The pattern of results was stronger for inexperienced students than experienced students. This is consistent with what is common in human-human interaction, where people may adjust the politeness level in a conversation based on how well they know the other party. This study contributes to the empirical base for developing pedagogical conversational agents that are socially sensitive. When the goal is to create conversations that are perceived as polite, direct commands and commands attributed to machines should be avoided. This guideline reflects an extension of politeness theory from HHI to HCI.

There have also been studies aimed at investigating the effects of agent apologies (one of the negative etiquette strategies in Brown and Levinson's model) in human-computer interaction (Tzeng 2004; Tzeng 2006). Tzeng (2004) found that computer apologetic messages helped create more desirable psychological experiences. Apologetic feedback made the subjects feel more respected, more comfortable, less mechanical and that the machine was more sensitive to a users' feelings compared with non-apologetic feedback. However, no significant difference was found between subjects in different apologetic feedback groups when evaluating user perceptions of their own performance. No significant differences were found among feedback treatments regarding the attractiveness and effectiveness of a computer game. In addition, the authors also found that subjects in the non-apologetic feedback groups attributed the outcome of guessing more to the game than to themselves, as compared with subjects in apologetic groups. This study also showed that applying etiquette strategies (in this case a negative strategy) can improve user psychological experiences, but not necessarily overall performance. Apology is the mostly widely used strategy to address a listener's negative face. Consequently, this is regarded as an important element in utterance messages that a service robot might present in a medicine delivery task.

As an extension to the above study, Tzeng (2006) investigated three computer politeness orientations along with three error message types and two frustration levels on subject performance in web searches. They found that high politeness-oriented subjects preferred apologetic messages significantly more than mechanical or joke messages and that subjects with low politeness orientations disliked the joke message most. These are very important observations as they reveal significant individual differences in politeness strategy perceptions. Overall, the participants found the apologetic message to be most friendly, whereas the joke message was the least friendly. This is especially true for those who preferred straightforward expressions of politeness. In addition, most of the subjects thought the mechanical message was the best indicator for error, and the joke was the worst one.

Wilkie, Jack et al. (2005) investigated whether the introduction of face-redressive expressions, based on Brown and Levinson's linguistic model of politeness, in the design of system-initiated proposals in automated telephone banking services, could mitigate the negative impact of task interruptions. Results showed that the presence of a proposal in the dialogue had a significant negative impact on the perceived overall service usability (it was perceived as more frustrating, less enjoyable to use, less efficient, etc.), while there were no differences in mean scores between the proposal groups. Consistent with the etiquette model, the negative face-redressive strategy was perceived as more formal, more polite and apologetic. The positive face redressive strategy was considered to be more patronizing and intrusive than the bald strategy. The bald strategy was considered to be shorter and less long-winded, and was favored by the participants. These results provided us with a validation of extension of the politeness model to HCI domain. Related to the current study, one interesting research question is whether such an extension is valid in the HRI domain.

Miller et al. (2004) conducted a field study with a senior focus group over a period from 4 to 6 months, in which medical reminder systems were put into use in a total of 11 living apartments. The automated medicine reminder system was capable of expressing five different reminding utterances according to Brown and Levinson's "face mitigation strategies", from most polite to least polite. The authors used them in the experiment to examine, among three types of users; including elders, engineers, and persons with nominal technical experiences, perceptions of system "politeness". User subjective feelings on overall system "appropriateness" were also collected as a response. Results showed that Brown and Levinson's model provided reasonable predictions of user perceived politeness except for "Off record" utterances. In human-human interaction, this type of strategy is highly indirect and context dependent, providing the speaker with plausible deniability for having made a request at all. However, such utterances might be difficult or even impossible for machines to produce. This utterance was perceived as "rude" in the experiment by Miller et al. (2004). Furthermore, in evaluating system appropriateness, it was found that the elderly did not see much difference between various utterance types, while younger groups followed similar patterns to each other. Overall, the "Off record" utterance was perceived as the most inappropriate because it did not provide a reminder message to take medicine. "Negative politeness" utterances were regarded as the most appropriate. In general, it was found from subject comments that the elderly were sometimes less comfortable with advanced technological systems or even tended to refuse to use the reminding systems. Thus, "politeness" and the etiquette behaviors may be important issues for elderly and could enhance their interaction experience. Furthermore, the authors argued that a purely polite medication reminder might not be the best one for the purpose of medication schedule compliance. They suggested that any politeness or etiquette-based systems should be highly flexible and adaptable to its individual user's expectations both initially and over time and context. This finding represents a major challenge from the perspective of designing for politeness and etiquette.

In summary, the above studies indicate that Brown and Levinson's model can be effectively used in HCI research. The politeness strategies might be effective from a HCI performance perspective. In general, a negative politeness strategy is preferred and can improve user psychological experiences. Previous studies used subjective ratings to measure user performance and therefore, different results were seen regarding the influence of etiquette on human performance (see Table 2-1). However, objective measures could be used to establish the relationship between etiquette strategy and human performance in future research. In this regard, the current study used objective measures to assess both robot task performance as well as user task performance, as influenced by robot etiquette strategies.

Table 2-1. The applications of Brown and Levinson's politeness model in HCI studies

	Strategy used	Main Findings
Wang, Johnson et al. (2005)	Direct (bald) and indirect (polite)	Polite tactic can affect students' motivational state and help students learn difficult concepts in interacting with pedagogical system.
Mayer, Johnson, et al. (2006)	Positive and negative	An extension of politeness theory from HHI to HCI was established.
Tzeng (2004)	Negative (apologetic and non-apologetic)	Negative politeness strategy can improve user psychological experiences, but not overall performance
Tzeng (2006)	Negative, positive and bald	Negative strategy is preferred and regarded as most friendly. Individual differences should be considered in the design etiquette.
Miller et al. (2004)	Negative, positive, bald, and "off-record"	Negative strategy is the most appropriate; the "off-record" strategy is the most inappropriate. User subjective perception of etiquette follows the patterns in Brown and Levinson's model
Wilkie, Jack et al. (2005)	Negative, positive, bald (control)	Negative face-redressive strategy was more formal, more polite and apologetic. Positive face redressive strategy was more patronizing and intrusive. Bald strategy was shorter and less long-winded. Validation of extension of politeness theory from HHI to HCI.

2.3.Studies of etiquette in human robot interaction

Walters et al. (2007) stated that HRI scenarios should be designed to seem as "natural" as in HHI scenarios. Unfortunately, robot behaviors investigated in prior research have been so trivial or obvious that few findings on HHI can be applied to such HRI. Related to this, Walter et al. observed that even if there was existing research of human-human interaction as a starting point for HRI design, it would not necessarily be the case that a robot would simply behave in the same way as a human in a comparable situation. In order to incorporate guidelines for HRI to be similar to HHI scenarios, further field research is required to increase our understanding on how social robots should behave under various scenarios. In this section, studies on HRI focusing on the social aspects of human interaction with intelligent service robots are reviewed. In addition, discussion on studies that have been conducted in the area of robot etiquette, with a focus on robots' respect of the user social space and shared environment, is provided. Finally, a list of differences in etiquette considerations between HCI and HRI are summarized.

2.3.1. Social aspects of interacting with intelligent service robots

Substantial research has been conducted to investigate the social aspects of interacting with intelligent service robots. For example, Dautenhahn, Woods et al. (2005) investigated human perceptions of, and attitudes towards, the idea of a robot companion in a home environment. They found that most of the participants were in favor of a robot companion (40%) and saw it serving an assistant, machine or servant, instead of a friend. Routine tasks, such as gardening, and vacuuming were expected for service robots more than caring tasks (e.g., child care). Humanlike communication was desirable, whereas humanlike behavior and appearance were considered less essential. In addition to that, Severinson-Eklundh et al. (2003) found that, in an office environment, robots' humanoid characteristics, the size, shape, color and the movements were important factors in contributing to its "personality". Their results revealed that addressing only the primary user in service robot applications was unsatisfactory. The setting, activities and social interactions of the group of people where the robot was to be used should also be considered by designers. Proper robot dialogue capability was suggested as a method to allow a system to successfully interact with group users.

There have also been studies conducted with social service robots in learning and gaming scenarios. For example, Kanda and Hirano (2004) examined the extent to which robots could form relationships with children and whether children might learn from robots in the same way they learn from other children. They found that long-term interaction with social robotic partners was valuable for language learning purposes. In addition to this, the authors found that a child's interest in robots decreased dramatically during the second week of exposure, which might have been due to limitations of the robot platform to be a true partner. Therefore, a humanoid robot was suggested because it might help in establishing more common social goals in the interaction.

However, the impact of such humanoid design features on the attraction and acceptance of robots could be opposite to what one expects, as suggested by Te Boekhorst et al. (2005). The researchers conducted an experiment to test the hypothesis that humans are more attentive towards robots if the robot appears to be "interested" in the people present. They found a significant interaction effect between the state of a camera and the state of a displayed pointer as an indicator of “attention by looking”. The children subjects paid more attention by looking at the robots when the pointer and camera were both static or moving) in the same direction as compared to when they pointed in different directions. Therefore, these findings indicate that humans may be less attentive to a robot’s request/task if there are conflicting effects in the robots’ behavior/interface (e.g., static pointer together with moving cameras).

Another relevant study was conducted to investigate user acceptance of robot assistance in the context of social, intelligent home environments (de Ruyter, Saini et al. 2005). The iCat, a robotic interface with the ability of simulating certain social intelligence, such as exhibiting facial expressions, vision and speech recognition was used in a Wizard of Oz experiment to assist participants in performing two tasks: program/record three broadcast shows with a DVD recorder, and participate in an online auction. Results showed that developing a home dialogue system with some social intelligence not only created a positive bias in user perceptions of technology, but also enhanced overall user acceptance. Participants could identify socially intelligent behaviors implemented in the system, and were more satisfied with the DVD recorder programming task under this condition. Participants who worked with a basic iCat (no social intelligence) were less inclined to continue working with the iCat at home. Although the interaction with the iCat was not the participants’ priority, beneficial effects were found for the whole home dialogue system when certain social intelligence was embedded. This research is directly relevant the present study, including whether robot etiquette/social intelligence has relevance to medicine delivery task performance.

In summary, this in-depth review of HRI studies provided some insights on how to conduct experiments on human interaction with a social service robot. In general, robot tasks and context characteristics, as well as robot appearance and communication abilities appear to influence HRI and should be considered in the design of future systems. Robot features should be compatible with each other and meet user expectancy. Conflicting interface features will result in decrease of user attention to robot requests (te Boekhorst, Walters et al. 2005), and might further affect user compliance with the robot. For design methods, a wizard-of-oz experiment was found to be appropriate to simulate desired robot intelligence. Video analysis was also helpful in capturing user natural reactions during interaction with a robot. A detailed summary of the studies, as well as their relevance to the present project is presented in Table 2-2.

Table 2-2. Summary of studies on human social robot interaction

	Scenario/ Robot Task	Design method	Results	Relevant issues be considered in this study
Dautenhahn, Woods et al. (2005)	Social companion in home environment	N/A	Routine tasks were expected from service robots vs. caring tasks. Humanlike communication was more desirable than humanlike behavior.	Robot tasks Robot communications
Te Boekhorst et al. (2005)	Game scenario	Video analysis	Humans were less attentive to robot's request if there were conflicting features in the robots' behavior/interface	Conflicting interface features
Severinson-Eklundh et al. (2003)	Servicing in office environment	Non-humanoid mobile robot. Additional character in the robot interface to simulate gestures.	Robot shape, size, color and movements were important factors contributing to its "personality". The setting, activities and social interactions of group of people where the robot is to be used should be considered by the designers.	Robot appearance Interaction settings
Kanda and Hirano (2004)	Learning scenario	Humanoid robot. Use of name tags for identification.	Long-term interaction with social robotic partners may be valuable. Humanoid robot may help to establish more common social goals in the interaction.	Humanoid feature
De Ruyter, Saini et al. (2005)	Intelligent home assistance	Wizard of Oz experiment to simulate robot intelligence.	Social intelligence in robotic systems may enhance overall user acceptance.	Social intelligence/adaptive assistance

2.3.2. Robot respect of human social space and shared environments

Walter et al. (2007) investigated how a robot should approach (frontal approach, left side approach, right approach, and rear approach) a human in a fetch and carry task in a home environment. The authors argued that in general, domestic or service robots should not simply move around and avoid people when carrying out tasks; they should also respect people's social spaces and shared workplace preferences. In an experiment, Walter et al. conducted a set of trials with four different robot approach scenarios in a converted seminar room. In the scenarios, subjects assumed different postures including: (1) sitting on a chair, (2) standing in an open space, (3) sitting at a table, and (4) standing against a wall. Subjective questionnaires were used to assess subject attitudes (such as comfort ratings and practicality ratings) and preferences for the different robot approach directions. The subjects were also asked to provide details about the reasons for preferring or not preferring particular robot approach directions after each trial.

The study results showed that most subjects do not like frontal approaches when seated, and even when seated behind a table. Instead they prefer to be approached from either left or right side, with a small overall preference for a right approach by the robot. Subjects do not like a robot to approach them from behind, preferring the robot to be in view even if it takes a physically non-optimal path. In the standing scenario, the rear central approach was rated as the least efficient and the frontal approach as the most efficient. However, in other scenarios, subjects did not identify a particular approach preference with regard to task efficiency. The authors stated that, in the standing condition, people found a robot directly approaching them to be less aggressive and invasive, particular if they were taller than the robot. A similar outcome was found between “standing in the middle of the room” condition and the “standing against a wall” condition, where the subjects rated the frontal approach as uncomfortable in the latter condition. This again related to the feeling of safety with the robot as it would be harder for subjects to escape in the standing against wall situation when the robot was approaching from the front.

In addition to these results, Walter et al. (2005) reported that the preferred robot approaching distance is within the expected ranges for comparable human-human social distance (0.45m to 3.6m). They also found that such preference would be mediated by human personality factors. For example, subjects who are more “Proactive” come less close to the robot. These findings are valuable references for the design of experiment with a mobile service robot.

2.3.3. Differences in human interaction with social robots and computers

Breazeal (2004) explored the topic of HRI from the perspective of designing sociable robots. She argued that robots that participate in a rich social exchange with users provide a number of unique advantages. First, people may find the interaction to be more enjoyable and comfortable. Second, communication is expected to require less training because humans are experts in social interaction. Third, teaching robots new tasks might be easier for users if the robots are capable of various forms of social learning such as imitation. To achieve these benefits, the author listed nine issues that should be considered and addressed in the robot design process. Breazeal posited that although the field of HCI might offer some insights for designing optimum interaction experiences for HRI, significant differences still exist, which make HRI a distinct area of inquiry. Several of the issues identified by Breazeal, which are directly relevant to this study, are highlighted as follows:

First, a robot is part of the physical environment and human interaction with social robots may change on a long-term basis. Initiation of interaction may be intentional or by chance. However, when interacting with software agents versus robots, one must go to a computer (or PDA) and such interaction normally constitutes a short-term relationship. When interacting with humans, robots also bring a set of particular affordances due to their physical embodiment (such as grippers for manipulation). These affordances may be richer with respect to conveying system functionality versus affordances of software agents or computers.

Second, there might be no observable interface that mediates interaction between the human and robots. Controls and interfaces, integrated in social robotic applications, may be more intuitive or implicit in form compared to HCI interfaces.

Third, robots work in a more dynamic and unpredictable environment and software agents tend to deal with more specialized tasks in a more restricted environment. In the future, robots will need to learn to "survive" in the real world. In other words, learning in human environments is a particular feature and challenge for effective human robot interaction. Robots will be expected to know what matters, what action to try, when to learn and who to learn from, how to correct errors and recognize success. These expectations are very different from those that users typically have for computers.

Related to this, Han et al. (2005) conducted an experiment to compare the effects of traditional media-assisted learning (books with an audio tape) and Web-based Instruction (web browser), with the effects of Home Robot-assisted learning. They focused on whether a home robot companion could make a difference in children's English learning abilities compared to other media. The robot they used was a commercial humanoid robot named *"IROBI"*, which is capable of presenting some facial expressions, such as happiness and calm. Results showed that children's concentration on material, and interests in English learning, were improved significantly in the robot-assisted condition. Moreover, the home robot was also superior in promoting and improving students' overall academic achievements. The authors, unfortunately, did not address in detail what particular robot features caused such differences. Table 2-3 summarizes and lists the features of the three media they used in Han et al. experiment. Three characteristics can be found that distinguish computers from traditional reading technologies including: (1) words for output, (2) interactivity, and (3) filling a role traditionally filled by a human (Nass and Moon 2000). These cues are incorporated into computers and may help encourage social responses. The same argument can be applied to social robots. Physical shape and humanoid features are additional cues that may trigger a human social response and expectations in HRI.

Table 2-3. Media features used in Han et al. (2005) adopted from Fong et al. (2001)

Media	Features
Book with audio tape	Recorded speech
Computer and web-browser	Text Words for output Interactivity
IROBI robot	Role Text on interface Facial Expression Movement Dynamic speech Humanoid body
HRI: Autonomous, three dimensional, dynamic user model/form, moveable HCI: Two dimensional, simple and static, vision and audio	

2.4. Human social responses and robot anthropomorphism

2.4.1. Human social responses to computers

Human social responses to computers and to anthropomorphic robots are important to understanding what features in design are most economical for influencing human etiquette expectations and behavior. Nass and his colleagues spent several years studying the effects of how human beings apply the social rules and stereotypes towards computers. In general, they found that humans *mindlessly* apply social rules and expectations. First, people rely on social categories when interacting with computers. For example, gender stereotypes have been found in human-social computer interaction (Nass, Moon et al. 1997). Results showed that humans found a computer tutor with a female voice to be significantly less friendly than a system with a male voice. The female-voice computer was rated to be less competent, but more informative about love and relationships, as compared to the male-voice system, which was rated to be more informative. Nass and his colleagues also conducted a study to determine whether people would rely on an arbitrarily assigned group (in group vs. out group), when interacting with computers (Nass, Fogg et al. 1996). Results showed that participants in "team" conditions are more likely to cooperate with computers, conform to a computer's suggestions, to assess the computer as being more intelligent, friendly and similar to themselves, compared with participants in a "non-team" group.

Second, humans apply over-learned social behaviors to computers; some examples include politeness and reciprocity. Nass et al. (1999) found that a human's evaluation of a computer was more positive if the same computer "asked" him to do a rating. Related to this, Fogg and Nass (1997) found that participants who worked with a helpful computer in a first task, and then returned to the same computer in Task 2, performed significantly more work for the computer in the second task, compared with a condition when the subjects used another computer for the second task.

A follow-up study to the Fogg and Nass research was conducted but with a focus on reciprocal self-disclosure in HCI (Moon 2000). There is substantial evidence, in human-human interaction scenarios that humans will engage in more intimate self-disclosure if they first become the recipients of such disclosures from their conversational partners. The results of the experiment provided evidence that self-disclosure tendencies in HCI are consistent with the norms of reciprocity. That is, responses in the reciprocity condition were higher in intimacy compared to the other two conditions.

The above findings suggested that humans will respond to computers and other technologies using the same social rules they use in human-human interaction. The extent to which social rules are applied depends on the number and strength of cues conveyed in the technology (Nass and Moon 2000). What has not been done up to this point is assessment of whether such phenomena extend to the area of HRI, where social service robots have the ability to express more and stronger social cues during interaction with humans. In particular, the current study explored whether humans can recognize the etiquette strategies robots use in communication, and how human perceptions of robot language etiquette may change between a humanoid robot and a neutral-looking robot.

2.4.2. Anthropomorphism and robot design

Anthropomorphism is defined in this project as “the intent and practice of designers to implement human-like features on a non-human object, such as computers and robots”. Here, anthropomorphism is basically considered as a design approach. However, from a practical standpoint, what matters most is the effect of such anthropomorphic features on people’s expectations of social etiquette in interacting with technology and their behavior towards a computer or robot. Robot anthropomorphism can be further classified into sub-categories based on two factors including physical anthropomorphism and behavioral anthropomorphism. The former factor refers to the extent to which the robot appearance resembles a human being, whereas the latter factor refers to the extent to which the actions of the robot mimic those of a human. In this project, focus was placed on the robot physical anthropomorphism. Behavioral factors such as expressions, movements, voice, were not manipulated in our experiment. A review of robot anthropomorphism, and how it drives human social responses can be found in (Zhang, Zhu et al. 2008).

Kanda, Miyashita et al. (2008) conducted an experiment to investigate human impressions of, and behavior towards, two robots with different degrees of anthropomorphism, where a human confederate was used as a baseline control. The Robovie, which has a head, two arms, a body, and a wheeled-type mobile base, represents a robot with a low degree of “humanness”. The Honda ASIMO is a biped humanoid robot with a higher degree of humanness. ASIMO hides all mechanical parts behind a white outer-shell. These robots were tested in a series of interaction scenarios during Kanda’s experiment. Results showed that the ASIMO received better subjective impressions than the Robovie or the human. However, different robot appearances did not result in a difference in participant verbal responses. Participants provided the same amount of information with identical politeness to ASIMO, Robovie, and a human confederate during the trials. It was also found that participants replied more rapidly to verbal greetings from the human confederate than the robots, and they replied more rapidly to an ASIMO greeting than to a Robovie greeting. When the robot asked the participants to look at a poster, it was found that participants looked most rapidly when Robovie pointed to it. In addition, participants tended to stand and walk closer to the ASIMO than to the Robovie robot. The authors argued that the white shading, rounded shape, biped walking mechanism, and hidden mechanical parts were critical design factors that caused the differences in participant perception of robot appearances/behaviors. In addition, distance and delay of response were also affected by robot appearance.

Related to the Kanda study, Hinds et al. (2004) examined the effects of robot physical appearance and status on how people work with robots in a collaborative task scenario. They were particularly interested in two responses, including: (1) human reliance on work partners; and (2) the extent to which people assume responsibility for tasks. The anthropomorphism of a work partner (a human confederate, human-like robot and machine-like robot) and the role of the work partner (subordinate, peer and supervisor) were manipulated with the human condition as a baseline. Results showed that humans relied more on the human confederate as compared with robotic partners. However, there was no evidence indicating that people will rely more on human-like robots as compared to machine-like robots. People did feel significantly less responsible for tasks when collaborating with a human-like robot partner versus a machine-like robot partner.

Another related study showed that anthropomorphic characteristics may contribute to robot acceptance and human compliance with robot behaviors (Goetz, Kiesler et al. 2003). Goetz and his colleagues demonstrated that a robot's appearance and behavior affect people's perceptions of the robot and their willingness to comply with its instructions. Participants preferred robots for jobs and would comply with them more when the robot's humanness matched the sociability required in those jobs.

In summary, prior research suggests that during tasks where human robot collaboration is needed, anthropomorphic features in robot design may contribute to overall task outcomes and interaction experiences for human users. The degree of humanness required depends on user expectations in the specific context and the degree of sociability required in the robot tasks. In general, a neutral-looking robot with some humanoid features has been found to be superior to a super humanoid robot. DiSalvo et al. (2002) found that the presence of a nose, a mouth, eyelids, skin, as well as the form of language are essential and greatly contribute to the perception of robot humanness. However, these features also suggest certain functional capabilities of robots. If robots do not meet human functional expectations, this can significantly degrade performance. Based on these findings, the present study sought to quantity the effect of robot anthropomorphism on user perceptions of (language) etiquette in interactive tasks.

3. RESEARCH OVERVIEW

In the context of human-social robot interaction, the major goal is to achieve both an effective and efficient collaborative work environment, or in other words, to help improve human performance as well as robot performance. Particularly in medication delivery tasks, patient responsiveness to and compliance with a robot request is important for the effectiveness of the HRI. Considering etiquette in robot interface design can help support achievement of this goal.

Figure 3-1 presents a research model for the present study. With respect to social medicine delivery robots, three general roles can be identified, including: (1) automated vehicles, (2) social actors, and (3) intelligent reminders. Navigation and localization are two required tasks performed when a robot serves as an automated vehicle. When the robot successfully finds a targeted location, its role changes to be a social actor with the aim of doing some simple greetings with a human, explanation of its intentions, and presenting the user with his/her medicine and dosage information. The robot should also remind patients to take the medicine if a failure to accept the medicine is detected. At this stage, the robot serves as an automated reminding system, and auditory reminding messages are presented to fulfill this role.

Three etiquette factors, corresponding to the three robot roles, can be identified that may affect the effectiveness of the HRI: (1) robot approach direction and stopping area (Woods, Walters et al. 2006; Walters, Dautenhahn et al. 2007); (2) humanoid cues/features, such as human physical traits; and (3) choice of linguistic strategy. These robot design factors may influence human perceptions of etiquette and further mediate user compliance with the robot. Each of these aspects of robot design needs to be investigated empirically.

The anticipated outcome of considering etiquette in robot design is a set of guidelines and a basis for future robot design. In this project, focus was placed on the factors of humanoid cues and robot linguistic strategies in perceptions of etiquette. In investigating robot linguistic strategy, Brown and Levinson's model of etiquette was applied to physical social entities (robots) instead of computer agents. The performance of each interactant (user and robot) and subjective perceptions of etiquette are critical to determining the HRI effectiveness.

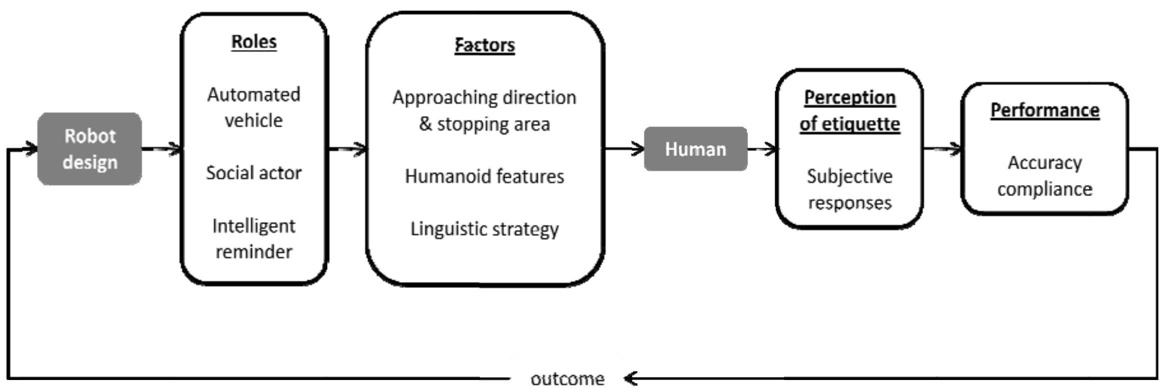

Figure 3-1. Considering etiquette in service robot design for a medicine delivery task

3.1.Research goal

The goal of this project was to apply classic etiquette strategies to social robot interface design, and assess how robot etiquette levels influence performance of humans and robots, in simulated medicine delivery tasks. The effects of robot physical appearance on user perceptions of linguistic politeness strategies were also assessed. To achieve this, two assistive robot prototypes with different levels of human physical likeness were prototyped.

There was also a control condition involving the use of a non-embodied agent (a laptop computer) for presenting information in the medicine delivery scenario. Participants listened to a medicine reminder message without interaction with a robot in the control condition. In addition, different types of reminding utterances were recorded and implemented according to Brown and Levinson's linguistic strategies for each robot platform. Questionnaires were used to measure a user's subjective feeling of perceived linguistic etiquette. These were developed based on the previous literature. A wizard-of-oz approach was taken to the experiment in which a human operator remotely controlled the robot prototype and played the recorded reminding messages. The experiment was used to quantity the relationship between robot linguistic etiquette and human perceptions of etiquette under the influence of robot physical anthropomorphism. Performance measures, including user compliance with robot requests and user task completion rate, were collected.

3.2.Research hypothesis

It was hypothesized that Brown and Levinson's politeness strategies in HHI could be extended to HRI in the medicine delivery and reminding scenario. In general, it was expected that users would be able to perceive differences in robot politeness based on the strategies used during the interaction. Such an extension of Brown and Levinson's model previously proved valid in HCI scenarios (Wilkie, Jack et al. 2005; Mayer, Johnson et al. 2006). Specific hypotheses on the various etiquette strategies included the following:

1) User overall perceived etiquette was expected to differ under different linguistic etiquette conditions. User overall perceived etiquette was expected to be higher under a high-etiquette condition versus a low-etiquette condition.
2) A negative etiquette strategy was expected to evoke more feelings of robot respect of a user's freedom to make decisions (e.g., to take or not to take the medicine) versus positive and bald etiquette strategies.
3) A positive etiquette strategy was expected to evoke more feelings that a user's value/wants/time were appreciated by the delivery robot in the social setting versus negative and bald etiquette strategies.

Second, as it has been found that good etiquette significantly enhances diagnosis performance in automation domains (Parasuraman and Miller 2004), it was hypothesized that applying politeness strategies would conform to social requirements during the service robot delivery task. Such design practice was expected to improve overall performance for both participants in their tasks:

1) The robot delivery task was assumed to be a disruption to any concurrent user task and applying etiquette strategies was expected to mediate adverse effects on user performance. User performance of a primary task was expected to increase when interacting with a robot with higher etiquette levels.
2) User compliance with robot task performance was expected to increase when interacting with robots with higher etiquette levels.
3) Consistent with the experiment results by Miller et al. (2004), it was hypothesized that negative politeness strategies could be most appropriate for medicine delivery and

reminding tasks. This strategy was expected to be more effective for user compliance with robot performance than positive politeness and bald strategies.

Goetz et al. (2003) found that a robot's appearance and behavior affects people's perceptions of the robot and their willingness to comply with instructions. Participants preferred robots for jobs and would comply with them more when the robot's humanness matched the sociability required in those jobs. Based on these findings, a third major hypothesis related to the influence of robot physical form/humanness on perceived etiquette and user compliance:

1) The presence of humanoid features in robots, in general, was expected to affect the perceived etiquette of linguistic strategies. User perception of robot etiquette was expected to differ between the humanoid robot and neutral-looking robot.
2) Humanoid physical features in robot interface design were also expected to mediate the effectiveness of etiquette strategies. In other words, an interaction effect between robot physical humanoid appearance and linguistic politeness strategies was expected. In terms of robot task performance (user compliance), a negative face strategy plus a robot with some physical humanness was expected to be better than all other combinations.

4. EXPERIMENT METHODOLOGY

4.1. Scenario and tasks

In the experiment test scenario, participants were instructed to act as patients following major medical surgery. They were provided with the following instructions: "Three days ago, you had appendicitis surgery and you are now going through the recovery phase in the hospital. This hospital recently purchased a nursing robot named "Pill-Me" to relieve the current workload of some nurses. The robot is adaptive and can recognize some simple speech and respond to you. The robot is also capable of exhibiting different levels of linguistic politeness under different situations. This morning, you are sitting in a patient common area watching a video on "how to play a Sudoku puzzle", and trying to complete one puzzle by yourself. Pill-Me comes in the room to provide you with a medical-related service for the first time. The robot will then leave the room when it "detects" task has been completed".

The experiment set up consisted of a simulated hospital room with the subject and the robot sharing the same environment. A Sudoku Puzzle game was used in the experiment as the primary task for the subject. Subjects were to complete the puzzle to the best of their ability within 15 minutes. There were two reasons for using the puzzle game in this experiment: (1) we wanted to measure user task performance as influenced by the robot etiquette strategy; and (2) the affect of etiquette strategy was expected to be more pronounced in situations when one interactant (the robot) interrupted another (the human) during the concurrent task performance. In other words, the experiment investigated whether the interruption to subject primary task performance amplified the affect of the robot behavior on perceived etiquette.

4.2.Task flow chart and robot route configuration

Figure 4-1 presents a flow chart of the tasks in the experiment. A trial began when the participants started watching the Sudoku video and playing the game. The robot began to move into the room at Time Point A. The robot wandered about the room for some time, and then left the room at Time Point B. This was the initial robot exposure for subjects. This set of events was used to provide subjects with an initial opportunity to view the robot without having to interact with it. This is similar to a real hospital situation in which a patient may simply see a robot in a common space before interacting with it in their room as part of a healthcare process. Such an initial exposure to the robot was also used to increase participant curiosity and state of arousal in advance of the medicine delivery. The robot was then sent back into the room at Time Point C and the robot attempted to provide the delivery service for subjects at Time Point D, with a maximum of three tries until it left the room. It was possible for subjects to ignore the robot and for the robot to fail in delivering the medicine. If a user did not respond to the robot after three attempts, this was considered a failure of user compliance, which will be discussed later. During the dwell time of the robot, the subject could either respond to the robot or concentrate on the game. A trial ended when the subject finished his/her primary task or after 15 minutes elapsed.

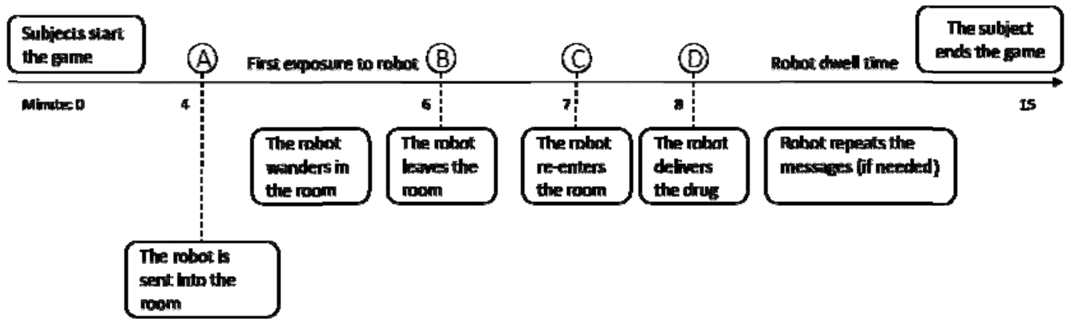

Figure 4-1. Task flow chart

Taking Walter et al. research as a basis, in the current study of etiquette in HRI, robot approach distance and direction were controlled and were constant in all trials. The service robot was programmed to navigate into the simulated patient room from the left side of the subject, and it stopped 1 meter away from the subject (see Figure 4-2). The subject remained seated when the robot came in to the room. However, they were allowed to stand up to interact with the robot if they wanted to. Negligible variances in subject perceptions of robot etiquette occurred based on these two factors.

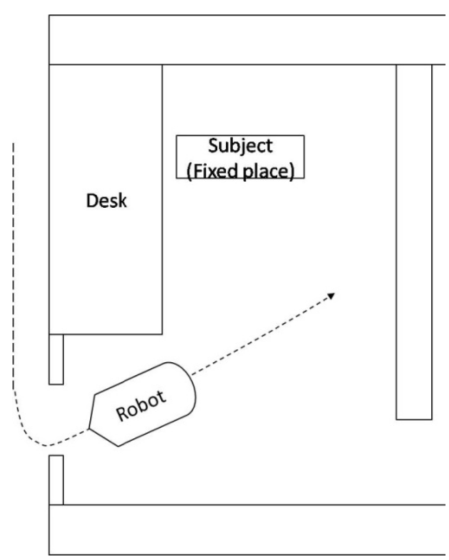

Figure 4-2. Robot route in the current project

4.3.Apparatus

A PeopleBot platform (Figure 4-3) was used in this research. The platform has autonomous navigation capability, including map-based route tracking and sensor-based obstacle avoidance. The robot was integrated with additional equipment to introduce different physical configurations for human-ness. A set of mini cameras and a face mask were used to present different human facial configurations. A tablet PC (HP tx2000) was mounted on top of the Peoplebot base and used as an on-board controller and to playback voice messages (WAV files) to users. Furthermore, a laptop PC was used for remote control of the robot, including navigation and playing utterances during the medicine delivery task. A digital video camera was used to record user interaction with the robot in each trial.

Figure 4-3. Basic PeopleBot platform

4.4. Participants

Although elderly persons (70 years and older) are expected to be the largest user group of service robots for medicine delivery or cognitive reminders in the future, the cost for recruiting senior participants and transporting equipment to hospitals can be very high (Zhang, Zhu et al. 2008). Therefore, this initial study of etiquette in robot design involved 32 university students. In order to successfully complete the tasks designed as part of the experiment, the subjects were required to have normal hearing capabilities (i.e., they did not require hearing aids on a daily basis). Minimum or no experience in playing a Sudoku puzzle game was also required for subject participation. The subjects indicated that they had not played a Sudoku puzzle within the past 6 months. They were only expected to have previously completed an easy-level Sudoku puzzle. Qualified participants were also required to have some experience with computers (familiarity with windows and how to use a mouse).

4.5. Variables

4.5.1. Independent variable

The independent variables in this study included (1) the robot physical appearance, and (2) robot language etiquette strategies. There were two levels of robot physical appearance including: (1) a humanoid robot; and a (2) neutral-looking robot. These two levels were separated in terms of the degree of humanness in robot design. To measure the degree of humanness, we counted the total number of humanoid features in the robot prototype design (based on DiSalvo et al. (2002) method). Prior research demonstrated that a robot head, face, mouth, and eyeballs are key features affecting perceived humanness (and influence human responses to robot requests). For the basic robot platform, two humanoid features were included: (1) a human-height body, and (2) a functional gripper (hand). By adding humanoid features to the robot face and body, we increased the degree of physical humanness in the robot prototype. The configurations of the two robot conditions are described in Table 4-1 and 4-2. The "Peoplebot + Costume + Mask" condition represented the robot with a higher degree of humanness, including a total of six humanoid features (the humanoid robot condition). In order to convey some of the identified characteristics of nurses, including authority, the appearance of cleanness, smoothness and professionalism, we donned the robot with a physician's coat. The doctor's coat was used instead of a nurse's scrub because the former provided for better cover of the mechanical parts of the original robot platform. The "Peoplebot + Mechanical eyes" condition represented a robot with lower degree of humanness, including a total of three humanoid features (the neutral-looking robot). In terms of transforming the mechanical looking robot to a humanoid looking one, the most salient aspects of design for promoting anthropomorphism were selected. The objective, however, was not to find an optimal form of humanoid robot for the medicine delivery task, or to replicate human appearance. Instead, the experiment was to test the effect of the salient humanoid features on human perceptions of robot etiquette.

Table 4-1. Physical humanness levels for two robot prototypes

Robot type	Peoplebot + Costume + Mask	Peoplebot+mechanical eyes
Features on the face and body	Eyes, mouth and costume (3)	Eyes (1)
Features of the platform	Human-height body Functional gripper (hand) (2)	
Total	3+2=5	1+2=3

Table 4-2. Two robot prototypes

Peoplebot + Costume + Mask	Peoplebot + Mechanical eyes
	

To manipulate etiquette strategies and robot utterances, four types of scripts were designed according to Brown and Levinson's politeness model. Each script included several phrases that the robot spoke to the subjects. For analysis purposes, we defined the total etiquette level of an utterance by counting the number of polite sentences in each script (counts were not doubled for repeated sentences). The scripts for each utterance are shown in Tables 4-3 to 4-6. Low etiquette utterances follow the bald strategy, defined by Brown and Levinson. The etiquette count for the least polite message (bald message) was -1 (one impolite phrase). Moderate etiquette utterances either represented positive or negative politeness strategies. High etiquette utterances were a combination of positive and negative politeness strategies with maximum counts of polite sentences.

Table 4-3. Low etiquette utterances

Utterance	Strategy used	Etiquette level
You have missed a dose of your medication. Accept and take your medicine now.	"Bald"	- 1 (One impolite highlighted phase.)

Table 4-4. Moderate etiquette utterance (Positive (P))

Utterance	Strategy used (positive)	Etiquette level
Hello, I know you must be busy, but your health is important. I have come to deliver medication to you. This should only take a second. Please accept and take your medicine now.	Claim common ground by seeking agreement/or by exaggeration of interest/approval/sympathy. Give (ask for) reasons from the observer. Attend to others. wants/needs/goals.	5 (Five polite highlighted phrases)

Table 4-5. Moderate etiquette utterance (Negative (N))

Utterance	Strategy used (negative)	Etiquette level
Excuse me, I am sorry to interrupt, but my nurse supervisor has indicated that you have not taken your medication scheduled for today. Here is the medication you need. I just want to ask you to confirm receipt of the medicine for me. If you have time now, could you please accept and take the medicine?	Give deference. Minimization of imposition Apology. Do not coerce the observer. Be pessimistic/use indirect requests.	5 (Five polite highlighted phrases)

Table 4-6. High etiquette utterance (P+N)

Utterance	Strategy used (P+N)	Etiquette level
Excuse me; I know you must be busy. I am sorry to interrupt, but your health is important. My nurse supervisor has indicated that you have not taken your medication scheduled for today. I have come to deliver the medication to you. This should only take a second. Here is the medication. I want to ask you if you would confirm receipt of the medicine for me. If you have time now, could you please accept and take the medicine?	Mixed strategies	9 (Nine polite highlighted phrases)

4.5.2. Dependent variable

The primary measure of perceived etiquette was a subjective rating across six major factors (see Appendix C).The participants were asked to evaluate robot language etiquette in terms of its relative disruptiveness, politeness, length of the message, usefulness/relevance, ease of understanding, and trustworthiness. After each robot trial, participants indicated the degree of robot etiquette along each of these six factors by ticking a box on a 7-point scale. The questionnaire was a revised version of the measure developed by Wilkie, Jack and Littlewood (2005).

Subject negative and positive ratings of robot etiquette were two additional dependent variables determined from the perceived etiquette questionnaire. For each robot utterance, a subject was asked to evaluate the degree to which the utterance evoked more feelings that the robot respected a user's freedom to make decisions (negative rating), as well as the degree that the utterance evoked more feelings that the user's value was appreciated by the robot (positive rating). The two responses were discrete in nature and ranged from 1 to 7 points on rating scales.

Some objective metrics were also used to evaluate social effectiveness. Many objective metrics, either from engineering, psychology, or sociology research, can be used to evaluate social "effectiveness" in the context of HRI (see Steinfeld et al., (2006) for an overview of HRI metrics). Here, we used two relevant measures as follows:

User compliance – This is the amount of cooperation a human gives to a robot, which may be critical for tasks in certain domains, such as health care. We used video recordings, similar to the approach used by Goetz and Kieler (2002), to assess compliance in the experiment tasks. User compliance was measured by how many times the robot repeated reminding utterances (with a maximum of three times). Compliance was also defined as whether the subject eventually responded to the robot request or not (success or failure).

Another measure of the degree of user compliance was response time to the robot request. This is a continuous metric and was used to indicate how reluctant a participant might have been to comply with the robot.

Total correct completion rate – This metric was used to determine how the robot etiquette strategy influenced user concurrent task performance. In this experiment, we counted the number of cells that a participant correctly completed in the Sudoku puzzle (Total correct completion). The total completion rate was measured by the proportion of correctly completed cells over the total number of completed cells.

4.6.Experiment design

In order to avoid the potential problem of treatment carryover effects in a within-subjects experiment design, a two factor between-subject design was used with level of etiquette strategy and level of robot physical humanness as independent variables. Robot physical humanness had two levels ("Peoplebot + mask + costume" and "Peoplebot + mechanical eyes"). The medium level humanoid robot was expected to provide fewer social cues and cause reduced expectations, as compared to a humanoid robot. Etiquette strategy had four levels including: (1) bald strategy, (2) positive etiquette strategy, (3) negative etiquette strategy, and (4) positive and negative etiquette strategy. Before formal test trials, participants completed a control condition trial in which they listened to medicine reminding utterances from a laptop computer. Overall, four data points were collected on each of the eight condition combinations.

Table 4-7. Experiment design

			Level of etiquette strategy			
			Bald	Positive (P) strategy	Negative (N) strategy	P+N
Robot	Level of physical human-ness	Basic robot platform + Mask + costume	n=4	n=4	n=4	n=4
		Basic robot platform + mechanic eyes	n=4	n=4	n=4	n=4

4.7.Subject instructions

Subjects were informed of the experiment scenario and tasks. They were also told that they would be awarded a gift certificate if their Sudoku performance was the best among all the subjects. A stopwatch was provided to the participants when they were playing the game to keep track of their performance time. The subjects were also informed that a robot would come into their room and provide them with a patient service. They were not told in advance that the robot would attempt to deliver medicine to them. Instead, the participants were told that the robot would give them some hints on solving the Sudoku puzzle (e.g., the robot might say *"Put 5 in row 2 column 3"*) after the robot finished his main task. Such a dependence of the primary task on the secondary task was created because the primary task response was used to assess the influence of robot etiquette on human task performance. The scenario ended after 15 minutes (a researcher came into the room and stopped the participant).

4.8.Experimental procedures

Subjects first read and signed an informed consent form and completed a background survey questionnaire (Appendix A). This questionnaire requested a subject's name, age, gender, level of familiarity with robots, prior experience with robots (if any), personal etiquette orientation, etc. After this, the experimenter presented a subject with several factors expected to influence perceptions of etiquette during interaction with other humans. Subjects were asked to compare pairs of factors and select the one that they thought was more important in evaluating language etiquette (Appendix B). Subsequently, the etiquette questionnaire used in the experiment was presented and introduced to the subject (see copy of form in Appendix C). In the baseline control condition testing, the medicine reminding utterances were played to participants (the same utterance as they heard from a robot in the formal test trials) using a laptop. We asked them to complete the perceived etiquette questionnaire based on the messages they heard. This ended the first part of the experiment.

In the second part, subjects were introduced the social service robots and their applications in hospitals. The experiment scenario and tasks were described in detail to the subjects. Subsequently, we began the formal robot testing. The subjects were required to complete the subjective perception of etiquette questionnaire again after interaction with the robot. A debriefing was provided at the end of the experiment. The total time for the experiment was about 1 hour.

5. DATA ANALYSIS AND RESULTS

5.1. Demographic statistics and data screening

The mean age of the participants (15 male and 17 female) was 24.3 years (SD=5.04). The majority of subjects (n=28) were not familiar with robots and had no direct exposure to robots previously, except from mass media. Two subjects said that they had interacted with robots, but they did not know, “how the robots worked”. Another two subjects said that they were familiar with the mechanisms of robots and knew what each part did.

Over 96% of the participants thought that social robots should be polite during their interaction with humans. Almost 40% of them expressed that robot speech should be kind and patient and that this was a key factor for robot etiquette (see Figure 5-1).

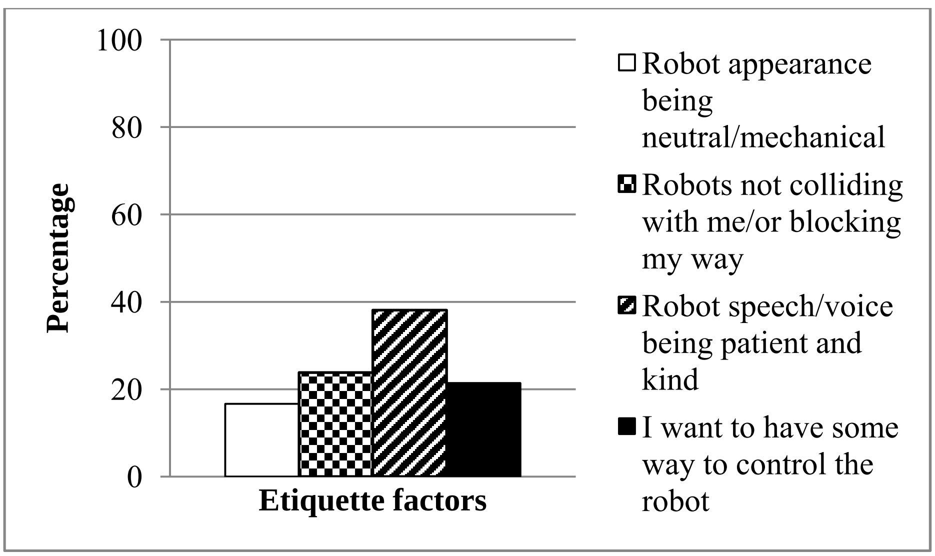

Figure 5-1. Etiquette factor and its percentage among subjects

With respect to subject etiquette orientation, it was found that most of the participants were either positive oriented (n=14) or negative oriented (n=14). Three subjects preferred the most straightforward strategy (bald). There was also one subject who chose the most conservative strategy and simply preferred not to do the FTA.

Finally, among the 32 subjects, 27 of them had no prior experience with Sudoku. Five of them said that they had tried an easy-level Sudoku puzzle before (more than 1 year ago), but none of them had ever completed a Sudoku puzzle. Other than that, all the subjects were regarded as amateur Sudoku players and their experience levels were the same after initial training instruction.

The subjective data from one subject was removed (Subject 15) from the data analysis as his ratings on perceived etiquette (PE) for the robot were equal to the ratings of PE for the control condition (=7, maximum). The subject did not follow the experiment instructions when filling out the questionnaire. However, the objective data for this subject (time to task completion, response time, etc.) was retained for statistical analyses. The descriptive statistics on all response measures for robot appearance and robot strategy are presented in Table 5-1 and 5-2, respectively.

Table 5-1. Descriptive statistic for different robot appearances

Robot appearance		CCR	RT	PE_control	PE_robot
Humanoid	Mean	0.26	19.64	5.20	4.78
	n	16	11	15	15
	Std. Deviation	0.19	15.81	1.14	1.18
Neutral	Mean	0.21	28.44	5.37	5.08
	n	16	9	16	16
	Std. Deviation	0.16	28.09	0.88	0.74
Total	Mean	0.24	23.60	5.29	4.93
	n	32	20	31	31
	Std. Deviation	0.17	22.00	0.99	0.97

CCR: correct completion rate.
RT: response time.
PE_control: perceived etiquette in control condition
PE_robot: perceived etiquette in robot condition

Table 5-2. Descriptive statistics for different etiquette strategy types

Etiquette strategy		CCR	RT	PE_control	PE_robot
B (bald)	Mean	0.24	25.25	4.29	4.67
	n	8	4	8	8
	Std. Deviation	0.12	30.80	0.95	1.29
N (negative)	Mean	0.32	17	6.35	5.66
	n	8	7	8	8
	Std. Deviation	0.27	12.99	0.44	0.52
P (positive)	Mean	0.15	23.20	5.09	4.39
	n	8	5	7	7
	Std. Deviation	0.11	21.25	0.54	0.83
P+N (positive and negative)	Mean	0.25	26.71	5.40	4.95
	n	8	4	8	8
	Std. Deviation	0.14	25.11	0.64	0.73
Total	Mean	0.24	23.60	5.29	4.93
	N	32	20	31	31
	Std. Deviation	0.17	22.00	0.99	0.97

5.2.Initial diagnosis of the experiment scenario

In order to test whether the experiment scenario influenced participant perceptions of etiquette in the reminding aspect of the medicine administration task (computer vs. robot), PE scores for the laptop condition were collected as a baseline for each subject. The mean PE score in the laptop condition was 5.35 (SD=0.9559), which was higher than for the robot conditions in general (Mean=5.04, SD=0.75). This indicated that the robot condition had an adverse effect on PE in the reminding task. Such adverse effect was partly due to the complexity and time constraints of the subject's primary task, as well as the disruptive nature of the robot delivery vs. the verbal reminders from the laptop. These conditions were expected to be different and this provided a basis (reason) for applying etiquette strategies to robot language in the delivery task.

5.3.Effect of etiquette strategy and robot appearance on PE

A two-way ANOVA model was used to test the effect of robot appearance and etiquette strategy on subject PE. (The GLM procedure in SAS was used to analyze the data.) The model residuals were checked for normality and constant variance across the levels of the fixed factors. Levene's test of equal variance showed that the error variances were not homogeneous among treatment groups (F(7,23)=5.441,P=.001). A Shapiro-Wilk test showed that the residuals were not normally distributed (W=0.89, P=0.0044). Common data transformation techniques were applied, however, they failed to satisfy the normality and equal variance assumptions. Due to the small sample size, Friedman's two-way nonparametric analysis was further conducted on this data set.

Results showed that subject PE ratings were significantly different among strategy types (χ_3^2 =93.41, P<0.0001). Post-hoc analyses using multiple comparisons revealed that the PE score in the negative condition was significantly higher than those in the other three conditions. The PE score for the negative condition was followed by the PE for the P+N condition. Interestingly, although the utterance in the bald condition had the least number of polite sentences (the sentences were scripted according to Brown and Levinson's original model), multiple comparison test results indicated that the PE score for this condition was not statistically different from that in P+N condition (see Table 5-3). That is, humans might prefer concise and clear requests during the interaction with social robots (even though it is not polite). This result also suggested that if a robot was designed to pretend to have the capability to conduct complicated social interaction, as humans do in real life, it might not always achieve the desired optimum PE.

The PE scores in the positive condition were actually the lowest among the four groups. There are two reasons for this: first, the language in the positive condition might not have been sensitive enough to subjects and hence, did not succeed in addressing, supporting and enhancing the positive face when applied to robots; second, it also possible that humans simply do not like the way (method) that a robot attends to their needs and wants (e.g., by way of exaggeration and claiming common ground). For example, one subject reported in the final survey that he thought the robot did not really know what he was doing and the robot's "interests" in his task were fake.

With respect to robot physical appearance, no significant difference in PE scores were observed (χ_1^2 =0.009, P=0.92). That is, robot appearance does not mediate PE in interaction while users are performing other cognitively complex tasks.

Table 5-3. Multiple comparisons of PE for different strategy types

Multiple comparisons of PE	Test statistic	P
B-N	$\chi_1^2 = 72$	<.0001
B-P	$\chi_1^2 = 4.84$	0.0277
B-P+N	$\chi_1^2 = 2$	0.1573
N-P	$\chi_1^2 = 43.62$	<.0001
N-P+N	$\chi_1^2 = 28.125$	<.0001
P-P+N	$\chi_1^2 = 19.384$	<.0001

5.4.User sensitivity to positive and negative etiquette strategies

Another two-way model was used to test whether a subject's etiquette rating was consistent with the original etiquette model, when it is applied to robot language. In particular, this analysis assessed whether a negative etiquette strategy would evoke more feelings that the robot respected a users' freedom to make decisions (Neg_rating) than the bald and positive strategies. It also served to determine whether a positive etiquette strategy would evoke more feelings that the user's value was appreciated by the delivery robot (Pos_rating) than the negative and bald strategies. Recall that the P+N strategy is a combination of positive and negative strategy, and is used to represent a higher level of etiquette. It was expected that the Neg_rating and Pos_rating in this condition would be the same as those in negative and positive conditions, respectively. Friedman's two-way nonparametric analysis was used to analyze the data. Robot appearance was treated as a blocking variable.

Results revealed that there was no significant difference in positive etiquette scores among the four strategy types (χ_3^2 =7.054, P=0.07). However, the negative ratings were significantly different among the four strategies types (χ_3^2 =112.07, P<0.0001). Post-hoc analysis using multiple comparisons showed that the negative rating in the negative strategy condition was significantly higher than that in the bald strategy condition (χ^2 =66.125, P<0.0001), as well as that in the positive strategy condition (χ^2 =43.61, P<.0001). This means that the negative ratings of subjects were sensitive to negative politeness strategies. Consistent with our hypothesis, no differences were found between the positive and bald strategy conditions (χ^2 =0.53, P=0.46), or between negative and P+N conditions (χ^2 =0.5, P=0.479), with respect to the negative rating (see Table 5-4). These results showed that participants were more sensitive to the negative strategy and thought that the robot respected their freedom to make decisions more in the negative condition as compared to bald and positive strategy conditions. This was consistent with the strategy definitions in original etiquette model. However, subjects could not interpret the positive strategy consistently with the original model. Therefore, compared with a positive etiquette strategy, the negative strategy is more effective for the purpose of conveying robot understanding of user needs and wants.

Table 5-4. Multiple comparisons of Neg_rating for different strategy types

Multiple comparisons of Neg_rating	Test statistic	P
B-N	$\chi_1^2 = 66.125$	<0.0001
B-P	$\chi_1^2 = 0.53$	0.46
B-P+N	$\chi_1^2 = 32$	<0.0001
N-P	$\chi_1^2 = 43.61$	<0.0001
N-P+N	$\chi_1^2 = 0.5$	0.479
P-P+N	$\chi_1^2 = 48.59$	<0.0001

5.5.Effect of etiquette strategy and robot appearance on user performance

A two-way ANOVA model was constructed to test whether the robot language strategy and physical appearance influenced subject primary task performance, in terms of correct completion rate (CCR). After an initial examination of the data, it was found that the CCR residuals did not satisfy the normality (Kolmogorov-Smirnov test, D=0.169, P=0.0201) or equal variance assumptions (Levene's test F(7,24)=3.146, P=0.017) of the ANOVA. Friedman's two-way nonparametric test was conducted to analyze the data. Results indicated that both robot appearance and strategy type had a significant influence on CCR ((χ^2 =10.56, P=0.00115 for appearance; χ^2 =50.88, P<0.0001 for strategy types). The post-hoc results using multiple comparisons revealed that the CCR in the humanoid robot condition was significantly higher than in the neutral-looking robot condition. The post-hoc comparisons also indicated that the CCRs in the bald, negative and P + N conditions were significantly higher than in the positive strategy condition (Table 5-5). In other words, the positive strategy was the most disruptive to the subject's primary task and led to the worst performance, compared to others.

In addition to assessing the effect of robot appearance and etiquette on user task performance, comparison was also made of the CCR among subject who complied with robot requests and those who did not. Figure 5-2 shows the extent of Sudoku puzzle completion for each group. As expected, it appeared that the subjects who engaged the robot, accepted the medicine and received a hint for primary task performance did perform better. This result validated the connection that was structured between the HRI and human task performance.

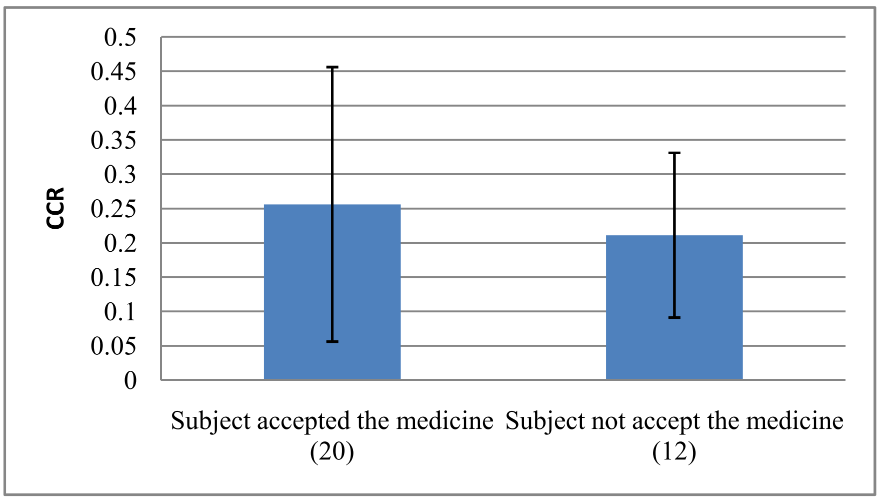

Figure 5-2. CCR for subjects who complied with robots and those that did not

Table 5-5. Multiple comparison of CCR for different etiquette strategies

Multiple comparisons of CCR	Test statistic	P
B-N	$\chi_1^2 = 0.5$	0.479
B-P	$\chi_1^2 = 32$	<0.0001
B-P+N	$\chi_1^2 = 0.001$	0.96
N-P	$\chi_1^2 = 18$	<0.0001
N-P+N	$\chi_1^2 = 2$	0.157
P-P+N	$\chi_1^2 = 36.125$	<0.0001

In order to test the interaction effect between strategy types and robot appearance on CCR, we recoded the interaction as a main effect in the statistical model (i.e., app*strat = 10*Appearance + Strategy). For example, the condition of humanoid robot with the bald strategy was coded as 1*10+1=11 (the coding rules for each of the feature types are shown in Table 5-6). A one-way non-parametric test was conducted on the interaction coded as a main effect. The result indicated that there was no significant interaction effect between strategy types and robot appearance (χ_7^2 =5.52, P=0.5958).

Table 5-6. Coding rules for feature types

Robot appearance	coding
Humanoid	1
Neutral	2
Etiquette strategy	**coding**
B	1
N	2
P	3
P+N	4

5.6. Effect of etiquette strategies and robot appearance on user compliance

The descriptive statistics on user compliance (UC) are provided in Table 5-7. Since the subject response to a robot request was measured as a success or failure, a binomial logistic regression analyses was conducted to test the influence of etiquette strategies and robot appearances on UC. Results indicated no significant effects of robot appearance (χ_1^2 =0.587, P=0.4436) or strategy types (χ_3^2 =2.7995, P=0.4236) on UC.

To further test a subject's reluctance in responding to the robot's request, the number of times the robot repeated the delivery message (utterance repetition) was analyzed (a discrete measure with counts of 1, 2 and 3; see Table 5-8 for details). Friedman's test results failed to reveal significant influences of appearance (χ_1^2 =2.25, P=0.1336) and strategy types (χ_3^2 =1.05, P=0.789) on the responses of utterance repetition. In general (across all subjects), the simple compliance measures did not appear to be sensitive to the robot appearance or linguistic strategy manipulations.

Table 5-7. Descriptive statistics of user compliance (UC)

User compliance	Frequency	Percent	Valid Percent	Cumulative Percent
Failure	12	37.5	37.5	37.5
Success	20	62.5	62.5	100.0
Total	32	100.0	100.0	

Table 5-8. Descriptive statistics of utterance repetition (UR)

Utterance repetition	Frequency	Percent	Cumulative Percent
1	14	43.8	43.8
2	4	12.5	56.3
3	14	43.8	100.0
Total	32	100.0	

5.7.Effect of etiquette strategies and robot appearance on response time

Only 20 subjects responded to the robot's request and response time (RT) was collected for these subjects. The RT was measured as the time from the robot's first request to the time when a subject accepted the medication. This was computed from the recorded video. This was a finer resolution measure of user compliance than the simple utterance repetition count. Friedman's test results indicated a significant influence of both the robot appearance (χ_1^2 =4.077, P=0.043) and strategy types (χ_3^2 =8.31, P=0.04) on user response time. Multiple comparisons for post-hoc analyses revealed that the subjects exposed to the humanoid robot condition responded significantly faster compared to subjects exposed to the neutral-looking robot condition. It was also observed that the RT in the positive, bald and P+N conditions were significantly higher than that in negative condition (see Table 5-9). That is, subjects also responded faster to negative politeness strategies in addition to rating them as posing higher PE.

Table 5-9. Multiple comparisons of response time for different etiquette strategies

Multiple comparisons of RT	Test statistic	P
B-N	$\chi_1^2 = 5.434$	0.0197
B-P	$\chi_1^2 = 2.308$	0.128
B-P+N	$\chi_1^2 = 0.333$	0.56
N-P	$\chi_1^2 = 4.289$	0.038
N-P+N	$\chi_1^2 = 3.24$	0.071
P-P+N	$\chi_1^2 = 0.92$	0.3367

To test the interaction effect on subject RT, a procedure similar to that used for the CCR analysis was followed. The interaction "app*strat" was recoded as a main effect. A one-way non-parametric (Kruskal-Wallis) test was conducted. The result failed to indicate a significant interaction effect between strategy types and robot appearance (χ_7^2 =4.8399, P=0.679) on RT.

5.8.Additional regression analysis on subject responses and objective responses

Additional regression models ($UC = \mu + PE\_robot + \varepsilon$; $UR, RT = \mu + PE\_robot + \varepsilon$) were structured to determine if perceptions of robot etiquette might drive user compliance with requests in the medicine delivery tasks. This analysis was aimed testing if there was any indirect influence of robot features and strategy types on user compliance and utterance repetition. First, a binomial logistic regression model for UC was analyzed. The result showed that there was no significant effect of PE_robot ($\chi_1^2 = 2.157, P = 0.1418$) on UC.

A PROC REG procedure was also conducted for UR and RT. Again, no effect was found for the PE_robot ratings (t (1, 29) =-1.17, P=0.25 for UR; t (1, 19) =0.82, P=0.422 for RT). These results indicated that perceived etiquette did not play a statistically reliable role in influencing user compliance or utterance repetition in this study. This is likely due to the fact that there was little variability in the UC according to the settings of the fixed effects.

6. DISCUSSION

6.1.Extension of etiquette model to HRI

In line with our expectation, user overall perceived etiquette for robot language was different under the different etiquette conditions. However, based on the current findings, the high-etiquette condition (P+N) did not result in higher PE, and the low-etiquette condition (bald) did not result in the lowest PE. A possible reason for this may be that people prefer a robot to be concise and clear in a delivery request. Lengthy and mixed-strategy utterances, although polite in nature, may cause confusion and should be implemented in robots with caution. Other than this, the PE score in the negative strategy condition was higher and the score was lower in the positive condition. That is, subjects may have interpreted and understood the robot negative utterance consistent with the original etiquette model. However, such consistency was not found for the positive utterance. In fact, no differences in positive etiquette scores were found among the four strategy types (this was contrary to the previous results found by Wilkie, Jack et al. (2005) in HCI domain). For this reason, the lowest PE score occurred in positive condition. Therefore, when an attempt was made to extend the Brown and Levinson etiquette model to the HRI domain, a positive strategy should be avoided, bald and P+N strategies should be applied with caution, and a negative strategy should be encouraged.

6.2.Robot etiquette and performance

Consistent with our hypotheses, user performance scores for the primary task were different under different etiquette conditions. Subject correct completion rates in the bald, negative and P + N conditions were significantly higher than in the positive condition. However, there was not enough evidence to indicate that the CCR would increase when interacting with a robot at higher etiquette levels. With respect to user compliance with robot requests, no significant differences were observed among the four etiquette strategy conditions. However, for the subjects who responded to the robot's request, they showed a faster response rate under negative etiquette conditions compared to other three conditions. This indicated that etiquette strategies do play a role in how reluctant a user may be in responding to a robot. These results are in line with findings from previous studies. For example, Miller et al. (2004) found that negative politeness strategies are the most appropriate for medicine delivery and reminding tasks. The authors also argued that a purely polite delivery message might not be the best one for the purpose of medication schedule compliance. This argument could help in explaining why high-etiquette utterances did not result in the least response time. In general, the results suggested that politeness strategies might be effective from a HRI performance perspective. Overall, the negative politeness strategy was preferred and can improve both the user performance and the robot performance in the task scenario.

Aside from the robot etiquette strategies investigated here, the gender of a voice in language presentation could also affect human perceptions of robot etiquette. Previous studies have found that humans apply etiquette rules specific to genders when interacting with computers (Nass, Moon et al., 1997). Other recent research (Mullennix et al., 2003) has demonstrated that a female voice is considered to be more persuasive than a male voice when presenting an argument using computer synthesized speech. In this study, however, we only used a male voice for presentation of the robot speech conditions. If a female voice had been used, results may not necessarily have been different among conditions; however, subjects might have been more responsive to the robots in general (i.e., a higher number of subjects complying with robot requests).

6.3.Effect of robot appearance on PE and performance

Partially consistent with our hypothesis, we found that simple changes in robot appearance do influence user performance and robot performance, but not PE. For example, subject correct completion rate for the primary task was higher in the humanoid robot condition, as compared with the neutral-looking robot condition. Subject response time to the robot request was also shorter in the humanoid robot condition. This finding was consistent with the findings from Kanda, Miyashita et al. (2008), in which they found that humans replied more rapidly to a humanoid robot greeting (ASIMO) than to a mechanical-looking robot greeting. Goetz (2003) also found that anthropomorphic characteristics may contribute to the elicitation of robot acceptance and human compliance with robot behaviors. A possible explanation for this result is from Nass's work. Nass said that humans will apply social rules and expectations to machines and robots with additional humanoid cues, and such over-learned behavior includes politeness and reciprocity. According to this theory, humans will adjust their interaction strategy during an experiment (or interaction scenario) to achieve mutual performance benefits when a humanoid robot comes to deliver medicine. Such a phenomenon is less likely to happen with a neutral-looking robot. Therefore, performance scores were lower in the neutral-looking robot condition. In addition to this, no interaction effect between robot physical appearance and linguistic politeness strategies were found for any performance responses. This implies that these two factors are not correlated and the robot appearance may not mediate the effectiveness of etiquette strategy.

7. CONCLUSION

Considering increases in health care service demands and the current shortage of nurses, service robots represent a technological solution for addressing routine patient service tasks. In such applications, there may be direct communication of patients with robots and robot language may play a critical role in the effectiveness of the interaction. This project assessed whether Brown and Levinson's etiquette model could be extended to the domain of HRI. Secondly, how such etiquette strategies influence both human performance and robot performance (user compliance), in a simulated medicine delivery task was evaluated. Finally, the effects of robot physical appearance on both subjective and objective responses were investigated. Findings indicated that not all strategies in the original etiquette model could be effectively extended to the domain of HRI. For example, in the positive strategy condition, subject PE did not follow the patterns prescribed by the model. The negative strategy was appeared to be the most effective across all responses. Therefore, caution should be taken when implementing these strategies in robots.

With respect to performance responses, a positive strategy degraded subject primary task performance; whereas a negative strategy decreased user response time to the robot request. Furthermore, humanoid robot features appear to provide additional social cues during the interaction. Such cues (or expectations) contribute to the elicitation robot acceptance and compliance with robot requests, as well as improvement in user primary task performance.

7.1.Limitations and future directions

Overall, the current study findings provide a basis for determining appropriate etiquette strategy types and robot appearances to promote better collaborative performance, as well as positive psychological responses in medicine delivery tasks to patients in hospitals. Some of the limitations of the study include: (1) subjects were not real patients and the sample size was small; (2) the lab experiment environment did not represent a real hospital environment and subject stress responses were likely different than would be expected in an actual hospital; and (3) each subject was only exposed to one test trial in the experiment; consequently, the reliability of individual subject ratings for a particular condition could not be tested. Furthermore, the manipulations made in robot appearance did not appear to be dramatic enough to influence subject expectations of robot capabilities. It is possible that a 3-dimensional robot head or active facial features would be more influential than a simple flat facial robot mask in generating false expectations during the test.

Additional research is needed on etiquette in human-social robot interaction, especially for health care industry applications, in order to confirm the current findings and extend them to a broader range of robot configurations. For example, there is a need to determine if the effectiveness of etiquette strategies differs among different cultures and age groups. Research also needs to explore how to quantitatively measure the effectiveness of a particular strategy, and how to use different combinations of mixed strategies. Beyond this, there is a need to determine if the effects of these strategies are task-dependent. For example, it would be interesting to investigate how the etiquette strategies influence FTA tasks with different severity levels. According to Miller's research in HCI, FTA severity could be quantitatively measured by social distance, power relations and act imposition. Taking this as a basis, different interaction scenarios could be developed for the medicine delivery task.

With respect to robot anthropomorphism, dynamic humanoid features, such as facial expressions, robot head movement and eye contact need to be investigated. Prior research has found that robot movements are important to human perceptions of anthropomorphism (Te Boekhorst et al. (2005)). In general, robot movements / gestures should be consistent to avoid false expectations for specific robot capabilities. Such results suggest that additional research should examine an off-record etiquette strategy to determine the influence of robot "body language" on human perceptions in medicine delivery tasks. To achieve this goal, simulated robot head avatars could be presented to users through a tablet PC on-board the robot, as an addition to the basic robot platform.

Effects of long-term interaction with a service robot, as well as team etiquette in HRI should also be a focus of future research. For example, it is still not clear whether a negative etiquette strategy can consistently improve human performance when a robot is interacting with a group of users over a long time period. Answering all these research questions may provide a good perspective for the growth of this new research area on etiquette in HRI.

8. **REFERENCES**

Breazeal, C. (2004). *Social interactions in HRI: the robot view*. Paper presented at the Systems, Man and Cybernetics, Part C, IEEE.

Breazeal, C., & Scassellati, B. (1999). *How to build robots that make friends and influence people.* Paper presented at the IEEE/RSJ International Conference.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*: Cambridge University Press.

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M. L., Koay, K. L., & Werry, I. (2005). What is a Robot Companion-Friend, Assistant or Butler? *Proceedings of IEEE IROS, Aug 2-6 2005.*

de Ruyter, B., Saini, P., Markopoulos, P., & van Breemen, A. (2005). Assessing the effects of building social intelligence in a robotic interface for the home. *Interacting with Computers, 17*(5), 522-541.

DiSalvo, C., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: the design and perception of humanoid robot heads. *Proceedings of the Conference on Designing Interactive Systems,* 2002

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: the design and perception of humanoid robot heads. *Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques*, 321-326.

Fogg, B. J., & Nass, C. (1997). Do users reciprocate to computers? *Proceedings of the CHI Conference* (Atlanta,GA). New York: Association of Computing Machinery.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems, 42*(3-4), 143-166.

Fong, T., Thorpe, C., & Baur, C. (2001). Collaboration, Dialogue, and Human-Robot Interaction*. Proceedings of the International Symposium on Robotics Research*,2001.

Goetz, J., & Kiesler, S. (2002). Cooperation with a robotic assistant. *Proceedings of the Human Factors in Computing Systems*, 578-579.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation*. Proceedings Workshop on Robot and Human Interactive Communication*: RO-MAN, 2003.

Goffman, E. (1967). Ritual Interaction: Essays on face-to-face behavior. *Pantheon, New York, NY.*

J. Han, M. Jo, S. Park, and S. Kim, "The Educational Use of Home Robots for Children," *Proceedings of IEEE International Workshop on Robot and Human Interactive Communication*, 2005, pp.378-383.

Hassmiller, S. B., & Cozine, M. (2006). Addressing The Nurse Shortage To Improve The Quality Of Patient Care. *Health Affairs, 25*(1), 268-274.

Hayes, C., Pande, A., & Miller, C. (2002). *Etiquette for human-Computer Work*. Massachusetts.

Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task. *Human-Computer Interaction, 19*(1&2), 151-181.

Janiszewski Goodin, H. (2003). The nursing shortage in the United States of America: an integrative review of the literature. *Journal of Advanced Nursing, 43*(4), 335-343.

Johansson, P., Oleni, M., & Fridlund, B. (2002). Patient satisfaction with nursing care in the context of health care: a literature study. *Scandinavian Journal of Caring Sciences, 16*(4), 337-344.

Kahn, M. W. (2008). Etiquette-Based Medicine. *New England Journal of Medicine, 358*(19), 1988.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human-Computer Interaction, 19*(1&2), 61-84.

Kanda, T., Miyashita, T., Osada, T., Haikawa, Y., & Ishiguro, H. (2008). Analysis of Humanoid Appearances in Human Robot Interaction. *Robotics, IEEE Transactions on [see also Robotics and Automation, IEEE Transactions on], 24*(3), 725-735.

Krishnamurthy, B., & Evans, J. (1992). HelpMate: A robotic courier for hospital use. *Systems, Man and Cybernetics, 1992., IEEE International Conference on*, 1630-1634.

Lyall, S. (2001). More Deaths In England Due to Error Retrieved Dec.20, 2001, from http://query.nytimes.com/gst/fullpage.html?sec=health&res=9B00E4D8133EF933A15751C1A9679C8B63&scp=1&sq=administering%20an%20incorrect%20amount%20of%20medicine%20&st=cse

Mayer, R. E., Johnson, W. L., Shaw, E., & Sandhu, S. (2006). Constructing computer-based tutors that are socially sensitive: Politeness in educational software. *International Journal of Human-Computer Studies, 64*(1), 36-42.

Miller, C. A. (2002). Definitions and dimensions of etiquette. *Etiquette for human-computer work*, 02-02.

Miller, C. A., & Funk, H. B. (2001). Associates with Etiquette: Meta-Communication to Make Human-Automation Interaction more Natural, Productive and Polite. *Proceedings of the 8th European Conference on Cognitive Science Approaches to Process Control*, 24-26.

Miller, C. A., Wu, P., & Chapman, M. (2004). The Role of "Etiquette" in an Automated Medication Reminder. *Proceedings of AAAI*.

Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research, 26*(4), 323-339.

Mullenix, J.W., Stern, S., Wilson, S., & Dyson, C-L (2003). Social perception of male and female computer synthesized speech. *Computers in Human Behavior*, 19(4), 407-424.

Murray, M. K. (2002). The Nursing Shortage: Past, Present, and Future. *JONA: The Journal of Nursing Administration, 32*(2), 79.

Nass, C. (2004). Etiquette equality: exhibitions and expectations of computer politeness (Vol. 47, pp. 35-37): ACM Press New York, NY, USA.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies, 45*(6), 669-678.

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Mindfulness Theory and Social Issues, 56*(1), 81-103.

Nass, C., Moon, Y., & Carney, P. (1999). Are respondents polite to computers? Social desirability and direct responses to computers. *Journal of Applied Social Psychology, 29*(5), 1093-1110.

Nass, C., Moon, Y., & Green, N. (1997). Are computers gender-neutral? Gender stereotypic responses to computers. *Journal of Applied Social Psychology, 27*(10), 864-876.

Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems (Vol. 47, pp. 51-55): ACM Press New York, NY, USA.

Park, H. K., Hong, H. S., Kwon, H. J., & Chung, M. J. (2001). A Nursing Robot System for The Elderly and The Disabled. *International Journal of Human-friendly Welfare Robotic Systems (HWRS), 2*(4), 11-16.

Parker, T. (2008). A Hollywood Family Takes on Medical Mistakes Retrieved Mar.17, 2008, from http://well.blogs.nytimes.com/2008/03/17/a-hollywood-family-takes-on-medical-mistakes/?scp=8&sq=administering%20medicine,%20error&st=cse

Perry, C. M. (1906). Nursing Ethics and Etiquette. *The American Journal of Nursing, 6*(7), 448-452.

Pollack, M. E., Engberg, S., Matthews, J. T., Thrun, S., Brown, L., Colbry, D., et al. (2002). Pearl: A mobile robotic assistant for the elderly. *AAAI Workshop on Automation as Eldercare, 2002*.

Preece, J. (2004). Etiquette online: from nice to necessary (Vol. 47, pp. 56-61): ACM Press New York, NY, USA.

Rogers, A., Hwang, W., Scott, L., Aiken, L., & Dinges, D. (2004). The Working Hours Of Hospital Staff Nurses And Patient Safety. *Health Affairs, 23*(4), 202-212.

Severinson-Eklundh, K., Green, A., & Hutenrauch, H. (2003). Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous Systems, 42*(3-4), 223-234.

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., et al. (2006). Common metrics for human-robot interaction. *ACM SIGCHI/SIGART Human-Robot Interaction*, 33-40.

te Boekhorst, R., Walters, M. L., Koay, K. L., Dautenhahn, K., & Nehaniv, C. L. (2005). A study of a single robot interacting with groups of children in a rotation game scenario. *Proc. of IEEE CIRA.*

Tzeng, J. Y. (2004). Toward a more civilized design: studying the effects of computers that apologize. *International Journal of Human-Computer Studies, 61*(3), 319-345.

Tzeng, J. Y. (2006). Matching users' diverse social scripts with resonating humanized features to create a polite interface. *International Journal of Human-Computer Studies, 64*(12), 1230-1242.

Walters, M.L, Dautenhahn, K., Te Boekhorst, R., Koay, K.L, Kaouri, C, Woods, S, Nehaniv, C, Lee D. and Werry, I. *The influence of subjects' personality traits on personal* *spatial zones in a human-robot interaction experiment*. Proceedings of 14th IEEE Int. Workshop on Robot & Human Communication (RO-MAN), (Nashville, USA, 2005), 347--352.

Walters, M. L., Dautenhahn, K., Woods, S. N., & Koay, K. L. (2007). Robotic etiquette: results from user studies involving a fetch and carry task (pp. 317-324): ACM Press New York, NY, USA.

Wang, N., Johnson, W. L., Rizzo, P., Shaw, E., & Mayer, R. E. (2005). Experimental evaluation of polite interaction tactics for pedagogical agents. *International Conference on Intelligent User Interfaces: Proceedings of the 10 th international conference on Intelligent User Interfaces, 10*(13), 12-19.

Whitworth, B. (2005). Polite Computing. *Behaviour & Information Technology, 24*(5), 353-363.

Wilkie, J., Jack, M. A., & Littlewood, P. J. (2005). System-initiated digressive proposals in automated human-computer telephone dialogues: the use of contrasting politeness strategies. *International Journal of Human-Computer Studies, 62*(1), 41-71.

Woods, S., Walters, M. L., Koay, K. L., & Dautenhahn, K. (2006). Comparing Human Robot Interaction Scenarios Using Live and Video Based Methods: Towards a Novel Methodological Approach. *Proceedings of the Ninth International Workshop on Advanced Motion Control (ACM 2006), Istanbul, March*, 27-29.

Zhang, T., Zhu, B., & Kaber, D. (2008). Anthropomorphism in Robots and Human Etiquette Expectations for Interaction. In C. Hayes & C. Miller (Eds.), *Human-Computer Etiquette*: Taylor and Francis.

Zhang, T., Zhu, B., Lee, L., & Kaber, D. (2008). *Service robot anthropomorphism and interface design for emotion in human-robot interaction.* Paper presented at the Automation Science and Engineering, 2008. CASE 2008. IEEE International Conference

APPENDICES

APPENDIX A

North Carolina State University

INFORMED CONSENT FORM for RESEARCH

Title of Study DESIGN OF ETIQUETTE FOR PATIENT ROBOT INTERACTION IN A MEDICINE DELIVERY TASK
Principal Investigator Biwen Zhu Faculty Sponsor (if applicable) Dr. David B. Kaber

---

What are some general things you should know about research studies?

You are being asked to take part in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate or to stop participating at any time. There is no direct benefit for you; however, you will have opportunity to interact with a real commercial assistive robot in this study. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask the researcher for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact the researcher(s) named above.

What is the purpose of this study?

The purpose of research studies is to examine the effect of applying classic linguistic etiquette strategies to social robot interface design, and to see how robot etiquette levels would influence the performance of both servers, in a simulated health care environment. We are also interested in studying the influences of robot's physical humanness on the users' perception of the linguistic politeness strategies.

What will happen if you take part in the study?

If you agree to participate in this study, you will be asked to sit beside a table, watch a video on "how to play a Sudoku puzzle game", and complete one by yourself in a simulated hospital environment. You need to complete the puzzle as much as possible within 15 minutes time and a gift certificate would be awarded if you complete the largest number of correct cells among all the participants. We will videotape your process of playing the game on the table. In the meanwhile, an assistive robot would come in and provide service for you. We will not tell you what kind of service the robot would provide for you. However, we do inform you that the robot would provide some hints for the puzzle game for you after "he" finishes "his" primary task. You have the right to decide whether or not to comply with the robot's request at specific time point. The experiment trial ends when you finish the game, which would take no more than 10 minutes. Together with the subject introduction and filling the survey, the total time for the experiment should be no more than 1 hour.

Risks

There are no potential risks involved in this study.

Benefits

There is no direct benefit for you; however, you will have opportunity to interact with a real commercial assistive robot in this study.

Confidentiality

The information in the study records will be kept strictly confidential. Data will be stored securely in a locked file cabinet in the Cognitive Ergonomics Lab in the Edward P. Fitts Department of Industrial and Systems Engineering, and will be destroyed at the end of the study. No reference will be made in oral or written reports which could link you to the study. You will NOT be asked to write your name on any study materials so that no one can match your identity to the answers that you provide.

Compensation

For participating in this study you will receive $10 compensation. If you withdraw from the study prior to its completion, you will receive $5 partial compensation system here. A $5 gift certificate would be awarded to the subject who finishes the task fastest.

What if you have questions about your rights as a research participant?

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919/515-4514), or Joe Rabiega, IRB Coordinator, Box 7514, NCSU Campus (919/515-7515).

Consent to participant

"I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may withdraw at any time."

Subject's signature____ **Date** _________________
Investigator's signature___________________________________ **Date** _________________

APPENDIX B

Background Information Survey

Welcome to our experiment. Let's start with some introductory questions to help us understand your background.

Name: ___________ Age: ____________Gender: _____________Major: _______________

Email: ____

1. How familiar are you with robots? [Circle the best answer]

   - o I have not had any direct exposure to robots except from mass media
   - o I have used them, but do not know how they work
   - o I am familiar with the parts of the robot and what each part does
   - o I could probably build a robot from parts

2. Your prior experience with robots [Circle all that apply]
   - o Work
   - o Toys
   - o In movies or books
   - o In TV shows
   - o In museums or in schools
   - o Robotic assistance
   - o Other:

3. Do you think it is important for social service robots to behave politely during the interaction with humans?
   - o Yes
   - o No

4. What are the factors you think that would influence users' perception of robot etiquette during the interaction? [Circle one that you think most important]
   - o Robot appearance being neutral/mechanical
   - o Robots not colliding on me/or block my way
   - o Robot speech/voice being patient and tender
   - o I want to have some way to control the robot
   - o Other

5. Suppose you laptop crashed and you have no idea how to fix it. You want to ask your friend Bob for help. However, you find Bob is currently busy on his own work in the office. What you would say to Bob in such situation. Please write down your speech down in the box with no more than 100 words. [Assumption: to fix the laptop is urgent and No.1 priority for you]

APPENDIX C

SUBJECTIVE COMPARISON OF FACTORS OF PERCEIVED ETIQUETTE

Indicate the characteristic of greater importance by circling its label on each line directly below:

Disruptiveness/Politeness

Disruptiveness/Length

Disruptiveness/Usefulness

Disruptiveness/Ease of understanding

Disruptiveness/Trustworthyness

Politeness/Length

Politeness /Usefulness

Politeness /Ease of understanding

Politeness /Trustworthyness

Length /Usefulness

Length /Ease of understanding

Length /Trustworthyness

Usefulness/Ease of understanding

Usefulness/Trustworthyness

Ease of understanding/Trustworthynes

Definition:

Disruptiveness – Whether the utterance is annoying, intrusive, distracting, or interrupts the user.

Politeness – Whether the utterance is polite, friendly, formal, apologetic, patronizing, manipulative, caring for individual needs.

Length – Length and long-windedness of utterance.

Usefulness – Whether the utterance is helpful, efficient, appropriate and relevant in content

Ease of understanding – Ease of understanding of messages.

Trustworthyness – Being confident in speech; willing to pursue the offer of service.

APPENDIX D

SUBJECTIVE PERCEPTION OF LINGUISTIC ETIQUETTE (CONTROL CONDITION)

Indicate the degree of etiquette of the recorded message for each of factors below by ticking the appropriate box along the scale.

The recorded message is not disruptive to my current job.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

The recorded message respects my needs/wants/time.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

The recorded message asserts minimum imposition.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

The recorded message is to the point.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

The recorded message is helpful, efficient, and relevant in content.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

The recorded message is easy to understand.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

The recorded message sounds confident throughout the service.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
□	□	□	□	□	□	□

APPENDIX E

SUBJECTIVE PERCEPTION OF LINGUISTIC ETIQUETTE (ROBOT CONDITION)

Indicate the degree of etiquette of the robot's utterance during the delivery and reminding task for each of factors below by ticking the appropriate box along the scale.

The robot speech is not disruptive to my current job.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

The robot speech respects my needs/wants/time.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

The robot speech asserts minimum imposition.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

The robot speech is to the point.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

The robot speech is helpful, efficient, and relevant in content.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

The robot speech is easy to understand.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

The robot speech sounds confident throughout the service.

Totally Agree	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Totally disagree
☐	☐	☐	☐	☐	☐	☐

APPENDIX F

Final survey

1. Do you feel the robot reminding utterance during the delivery service polite? If yes, write down the specific sentences; if not, how should it be improved?

2. Does the robot appearance look more or less like a real human being? What physical features you think should be added for a humanoid social service robot?

3. Would your perception of linguistic politeness to the specific reminding message change when a robot delivers that message to you, compared with listening to it from the laptop?

4. Would your perception of etiquette to the robot utterance change if the robot becomes more and more physically humanoid? If yes, how would it change? If not, explain briefly.