5-2015

Examining the Effects of Robot-Enacted Guilt Appeals in a Human-Robot Negotiation

Brett Stoll
Western Michigan University

EXAMINING THE EFFECTS OF ROBOT-ENACTED GUILT APPEALS IN A HUMAN-ROBOT NEGOTIATION

by

Brett Stoll

A thesis submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Master of Arts
School of Communication
Western Michigan University
May 2015

Thesis Committee:

Chad Edwards, Ph.D., Chair
Autumn Edwards, Ph.D.
Kathleen Propp, Ph.D.

EXAMINING THE EFFECTS OF ROBOT-ENACTED GUILT APPEALS IN A HUMAN-ROBOT NEGOTIATION

Brett Stoll, M.A.

Western Michigan University, 2015

Social robotics is a quickly evolving and expanding field in which significant contributions may be made by the communication discipline. Prior research has demonstrated the successful employment of robots throughout varying contexts such as work team decision-making, education, and healthcare. The purpose of this study is to expand upon existing research and generate an understanding of how robots may be used in competitive communication environments. The study highlights face negotiation theory (FNT) and the computers are social actors (CASA) paradigm to frame predictions and understanding of how humans interact with robots in a negotiation context. The researcher uses a 2 x 2 experimental method to examine how variables of guilt persuasion and robot agency influence human concession in a human-robot negotiation game. Statistical analysis of overall participant concession during negotiation indicates that there is no significant difference in how humans negotiate with a robot based on the robot's enactment of guilt appeals or whether the robot is positioned as a principal or agent negotiator. Further research is suggested to examine to what extent face might play a role in human-robot interaction, both in cooperative and competitive contexts.

TABLE OF CONTENTS

Table of Contents -- Continued

LIST OF TABLES

CHAPTER I

INTRODUCTION

Most people familiar with negotiation can relate to the classic car sale scenario. You are sitting at a desk across from the well-dressed salesperson as she presents the details of your new car purchase. Her goal – to sell you the four-wheel drive, limited edition behemoth with leather packaging, navigation, seating for 13, and an exclusive, chrome-dragon hood ornament. Your goal – to buy something practical to haul you and your dog around the city and at a price that does not instill near immediate buyer's remorse or eventual bankruptcy. After much haggling over ticket price, financing options, and included maintenance, an agreement is reached, and hopefully, you and the saleswoman both go home feeling satisfied.

Negotiating, whether for a new car, whose turn it is to do the dishes, or the details of a multi-billion dollar acquisition, can be a tricky and sometimes daunting prospect, especially for those of us with little skill or experience in such scenarios. No negotiation is the same. Not only do the faces change, but also do the expectations, attitudes, values, and desires brought to the table by each person. However, there is the catch – person implies *human* communication. The future of negotiation may further evolve as we are pitted against negotiators made not of flesh, blood, and emotions, but ones instead comprised of plastic, metal, circuits, and mathematical logic. We may soon realize the employment of robots as negotiation adversaries.

As far-fetched as negotiating with a robot may seem, technology may not be far from making this a reality. Already, computers and robots interact as companions

(Lehman, Iacono, Dautenhaun, Marti, & Robins, 2014) as well as partners and teammates in decision-making processes and the completion of other complex tasks (Kaupp, Makarenko, & Durrant-Whyte, 2010). Just as these entities may work with us, they may too work against us. Of course, this is not meant to inspire thoughts of judgment day and the eventual enslavement of humans by robots. Instead, it is merely acknowledging the role that more sophisticated technology may play in reducing the level of human involvement in certain contexts. Using computers and robots to fulfill surrogate human roles is not a new notion. Many of the robots employed today have very little interaction with the majority of humans. Others, such as automated phone systems and automated teller machines (ATM), currently address many of our day-to-day concerns, fulfilling more of a social interaction and human service function.

The implications for social robots, robots capable of interacting with humans based upon a set of social rules, span disciplines and have the potential to positively affect human learning (Robins, Dautenhahn, Boekhurst, & Billard, 2005; Tanaka, Cicourel, & Movellan, 2007), health (Breazeal, 2011; Scasesellati, 2007), task management (Breazeal, Kidd, Thomaz, Hoffman, & Berlin, 2005; Goetz, Kiesler, & Powers, 2003; Hinds, Roberts, & Jones, 2004), and overall social wellbeing (Hutson, Lim, Bentley, Bianchi-Berthouze, & Bowling, 2011; Kahn, Freier, Friedman, Severson, & Feldman, 2004). For example, Broekens, Heerink, and Rosendal (2009) highlight abundant evidence that social robots play a critical role in the care of elderly persons in Japan. They have been demonstrated to help improve mood, decrease loneliness, and encourage social connectivity. However, the authors remark that the studies reviewed tended to have research designs lacking robustness and were largely conducted in Japan, a culture at the forefront of humanoid social robot development and incorporation throughout society. This allows some, but not extensive, application to other cultures that are relative strangers to these forms of technology. In education, robots have been found to be useful tools in tutoring Japanese elementary students (Kanda, Hirano, Eaton, & Ishiguro, 2004) as they helped to establish common ground with students through social interaction and improved their retention of English. Kanda and Ishiguro (2005) additionally found evidence that robots not only improve learning of foreign language, but also demonstrate effectiveness at establishing longitudinal relationships with students. They argue that further development of social robotics has significant implications for providing student playmates and highly effective tutors. These few studies alone demonstrate the possibilities of widespread implementation of robots in a variety of contexts.

Robots may enact a social component of human interaction, but they are obviously not human. Robots are not prone to anxiety, jealousy, or hasty generalizations fueled by anger. Even the most task-oriented sociopath is still human and, as a result, is susceptible to emotionally driven motives and behaviors. Herein lies the benefit of moving robots from a strictly assistive function to a more competitive role. Naturally, conflict is emotionally charged and peaks both positive and negative arousal in people (Bodtker & Jameson, 2001). Introducing a robot to the interaction as primary negotiator would serve to reduce the emotional influence that occurs in negotiation. Although a programmer may have an ego, a robot does not. Rather, a robot functions on specific directives to achieve a favorable conclusion according to a cost-benefit analysis assigned by the programmer. However, it may be less about the effectiveness of the robot's reasoning abilities, and more about the effectiveness of the robot in interactions with humans. How does a human respond to a robot that has been positioned as an opponent? The concept of robots have existed long enough that they are a part of most peoples' perceptual sets, but how does the relative novelty and uncertainty of an adversarial robot affect negotiation processes and outcomes? Although robots may be immune to emotion, is it possible for a robot to strategically trigger emotional responses such as guilt in a human? Additionally, to what extent will the effectiveness of guilt appeals differ if participants feel guilt for the robot or toward another human controlling the robot?

The purpose of this study is to further the discussion of human-robot negotiation by examining the employment of guilt appeals by a robot and the role of perceived motive during negotiation. Specifically, this paper will investigate whether a robot can effectively employ emotional appeals regarding fairness in negotiation, seeking to elicit feelings of guilt and influence the behavior of a human negotiator. Additionally, this paper will address how motive assigned to a robot (whether the robot is negotiating on behalf of itself or on behalf of another human) affects how a human opponent perceives the robot and how the human responds to offers in a negotiation.

In order to address the research questions posed above, a literature review follows, examining the relevant research pertaining to Face Negotiation Theory (FNT; Ting-Toomey, 1988). FNT will be used as a theoretical framework to discuss how humans respond to one another during negotiation as well as to predict how humans may respond to a social robot as well. Following a discussion of FNT, this paper will examine the similarities and differences in Human-Human Interaction (HHI) and Human-Robot Interaction (HRI). Zhao (2006) and others argue that humans may interact with humanoid robots similarly to how they interact with other humans, applying the same social rules and behaviors. This point is further highlighted by the Computers are Social Actors (CASA) paradigm (Nass, Steuer, Tauber, & Reeder, 1993), included as a framework for further comparing HHI and HRI.

CHAPTER II

LITERATURE REVIEW

Conflict and Negotiation: Framing a Human-Robot Scenario

Negotiation is a widely studied topic in a variety of contexts given its application in nearly every aspect of communication. People use negotiation as a tool to navigate conflict, defined simply by Deutsch (1973) as the occurrence of incompatible activities. With such a simple definition of conflict, one may be able to identify a swath of events that took place just yesterday at work that match this description. Perhaps you and a colleague needed to use the copier at the same time, you had two meetings scheduled to occur at the same time, or maybe a team member confronted you with a conflicting, "better" plan for how best to manage a project on which you were assigned to work together. Regardless of the reason, in order to resolve the conflict, negotiation in some form must take place. Bazerman and Neale (1983) define negotiation as a joint decision-making process in which the allocation of scarce resources is determined by two or more parties. Given that the literature on negotiation is vast, the following section will highlight the literature focusing on Face Negotiation Theory and its implications for how individuals incorporate varying styles for managing conflict. This is later followed by a discussion of the relevant literature regarding human-robot communication and implications for negotiation.

Face Negotiation Theory

Face Negotiation Theory (FNT) claims that face is an explanatory mechanism for one's chosen conflict styles (Ting-Toomey, 1988, 2005). Face is the favorable image that an individual claims and seeks to protect in social and relational networks through behaviors known as facework. Important aspects of face include an individual's desire to be (a) viewed as competent and likeable by others to maintain interdependence (positive face) and (b) to be autonomous from others such that they can achieve individual goals free from interference (negative face; Ting-Toomey et al., 1991). Facework can be used as a means to protect one's personal face or to support or attack the face of another party (Ting-Toomey, 1988, 2005). At its core, FNT posits that communication is an ongoing process of facework as individuals are in a perpetual state of negotiating self and other-face. Although all humans negotiate face, various cultures engage in facework differently (Oetzel, et al., 2001; Ting-Toomey, et al., 2000; Trubisky, Ting-Toomey, & Lin, 1991). The purpose of the current research is to test whether humans feel the need to protect self-face or the face of others by conceding to guilt appeals enacted by a robot. In other words, will participants "feel badly" about treating a robot unfairly such that it provokes a need to engage in facework with the robot and influences how they negotiate?

Ultimately, the manner in which individuals engage facework is dependent upon one's self-construal and face concerns (Oetzel & Ting-Toomey, 2003). Self-construal is defined by Markus and Kitayama (1991) as one's self-perception of self-image traits, which can be either independent or interdependent. In other words, whether the individual places emphasis on his individuality and uniqueness or his connectedness and relationships will influence how he engages in facework. These two elements of individual self-construal (independent vs. interdependent) are often associated with their larger cultural dimensions, individualism and collectivism respectively (Ting-Toomey, Oetzel, & Yee-Jung, 2001). In conjunction with face concerns, the focus of attention, energy, or resources in a conflict is a factor (Ting-Toomey & Kurogi, 1998), which can be self, other, or mutually directed, that influences an individual's implementation of certain conflict styles.

Prior research (Zhang, Ting-Toomey, & Oetzel, 2014; Oetzel & Ting-Toomey, 2003) has highlighted the difficulty of operationalizing mutual concern beyond a theoretical discussion and as such has focused primarily on measuring self and other concern. As a result, these two dimensions of concern (concern for self, concern for other) were incorporated in a conflict styles model that assigns five conflict styles according to variations in one's concern for self and concern for other (Pruitt & Carnevale, 1993; Rahim, 1983; Thomas & Kilmann, 1974). The five conflict styles, integrating, competing, compromising, avoiding, and obliging, are structured in the model as follows. Understanding these conflict styles is important, as it is often referenced in conjunction with FNT research (Oetzel, 1998; Oetzel, Meares, Myers, & Lara, 2003; Ting-Toomey, et al., 1991), and it foregrounds the current study, which seeks to determine what conflict style behaviors a human is likely to enact in negotiation with a robot. Integrating is associated with high concern for self and other and is often termed the "win-win" model as parties seek to provide satisfactory outcomes for themselves and others. Competing incorporates high concern for self and low concern for other, often associated with someone willing to impress their demands to achieve an individual goal with little to no care of the opposition's outcomes. Compromising is comprised of moderate concern for self and other, involving concessions on behalf of both parties in order to reach an agreement and sometimes referred to as "We both win and lose". Obliging indicates low concern for self and high concern for others; the individual enacting this conflict style typically gives in to the demands of opposition in order to maintain harmony. Lastly, avoiding is associated with low concern for both self and other, where the individual seeks to simply withdraw from the conflict. With an understanding of FNT and the prominent incorporation of conflict styles in the FNT literature, a more focused understanding of emotions in conflict is needed.

FNT has been a prominent theory in conflict research, but until recently, the role of emotions in FNT has been largely unaddressed (Zhang, Ting-Toomey, & Oetzel, 2014). Prior research has indicated that emotion plays a role in the resulting conflict process and behaviors (Allred, Mallozzi, Matsui, & Raia, 1997; Bodtker & Jameson, 2001; Van Kleef, De Dreu, & Manstead, 2006). Zhang and colleagues (2014) examined how emotions of anger, compassion, and guilt had an effect on the specific conflict styles chosen by individuals during conflict. Results indicated that anger was positively associated with a competing style, compassion was positively related with integrating, compromising, and obliging styles in both China and the U.S., and guilt was positively associated with an obliging conflict style for U.S. participants and avoiding for Chinese participants.

Of particular interest to this study are guilt and its effects in conflict negotiation. Guilt has been defined as a social phenomenon (Baumeister, Stillwell, & Heatherton, 1994) in which the negative evaluation of one's personal behavior resulting from the perceived harm inflicted upon a relational partner inspires feelings of remorse and possible recourse to remedy the harm (Covert, Tangney, Maddux, & Heleno, 2003; Van Kleef, et al., 2006). Feelings of guilt elicited in negotiation have been associated with increased empathy and perspective taking (Leith & Baumeister, 1998), as well as behaviors of compliance and efforts to either prevent transgressions (Baumeister et al., 1994) or counter the harmful actions that have already occurred (Covert, et al. 2003). In this respect, guilt is an effective tool in conflict for seeking appeasement (Van Kleef et al., 2006). Ketelaar and Au (2003) demonstrated that increased feelings of guilt were positively associated with one's willingness to cooperate in human-human negotiation games. The researchers tested this by having participants play two difference negotiation games (Ultimatum Game & Prisoner's Dilemma Game) and experimentally manipulating the feelings of guilt. Other researchers have determined similar findings (Broekens, Jonker, & Meyer, 2010; Mandel & Dhami, 2005), as well as asserting that players and negotiators make decisions based on the objective of avoiding guilt (Battigalli & Dufwenberg, 2007; Dufwenberg, 2002). What remains unknown is if these feelings of guilt can manifest in human-robot negotiations and replicate similar results of participant obliging and conceding as examined in prior research (Zhang et al., 2014).

How Do We Communicate with Computerized Entities?

With the previous discussion of negotiation in mind, the following section reviews current research regarding how humans interact with computerized entities, any non-human object that operates as a function of computer programming. The CASA paradigm (Nass et al., 1993) foregrounds this discussion, demonstrating that humans treat computers in a similar manner to how they treat other humans, even in situations with limited social cues. The CASA paradigm is then expanded to incorporate specifically human-robot interaction, which eventually seeks to draw connections between human-human and human-robot negotiation.

Computers as Social Actors

The CASA paradigm (Nass et al., 1993) suggests that humans will communicate with computers similarly to how they would communicate with other humans. Further, one's tendency to anthropomorphize a computer exists even in the presence of minimal social cues and despite the user's explicit awareness and acknowledgement that the computer is neither human nor deserving of human-like treatment. Nass and colleagues (1993) term this process as *ethopoeia*. Since its inception, the simplicity of CASA has theoretically framed numerous research studies seeking to affirm that humans treat computers and humans similarly. The process of applying CASA is simple as it involves identifying a study of interpersonal human-human interaction and replicating the study with a computer substituting for one of the humans. According to CASA, such research methods should expect to replicate similar results as well.

Study replications such as these have been conducted many times by Nass and others. Using CASA as a framework, researchers have determined that individuals identify and assign human personalities such as dominance and submissiveness to computers based on their interaction in the classic Desert Survival Scenario (Nass, Moon, Fogg, Reeves, & Dryer, 1995), and that people make mindless attributions such as gender (Lee, Nass, & Brave, 2000), ethnicity (Nass, Moon, & Green, 1997), politeness, reciprocity (Katagiri, Nass, & Takeuchi, 2001; Nass, Moon, & Carney, 1999), and expertise (Nass & Moon, 2000; Nass, Reeves, & Leshner, 1996) to computers as well.

In an effort to understand why the CASA effect exists, Nass and Moon (2000) discussed a number of potential reasons, including social cues lending to anthropomorphic application, overlearning, and mindless attribution. Ultimately, the authors argued that the latter was the most likely of reasons, but that there may be varying degrees to which mindless attribution comes into play based upon the primitiveness of responses or the frequency of use for a particular social rule (Nass & Moon, 2000). Additional studies have been conducted to examine the difference between perceptions elicited by a computer and those elicited by a human. In these studies, all factors of the scenario are controlled with the exception of a single condition. Participants are informed that they are either communicating with a human via CMC or communicating with the computer. For example, Fogg and Nass (1997) examined the role that flattery played in HCI. Results demonstrated that flattery communicated by a computer, regardless of its genuineness, produced similar effects as flattery communicated by another human. Even features such as non-human, computerized voices were deemed capable of manifesting personalities that encouraged similar attributions and treatments of computers as demonstrated to humans with similar personalities (Nass & Lee, 2000; Nass, Moon, Fogg, Reeves, & Dryer, 1995). A computer's ability to elicit similar perceptions and behaviors to that of a human highlights their effectiveness as teammates (Nass, Fogg, & Moon, 1996) and other potential social applications such as companions (Tung, Sato, Deng, & Lin, 2009). Despite substantial evidence supporting the CASA paradigm, the question remains as to whether this theoretical framework applies to HRI as well as it applies to HCI. The following section seeks to establish this connection.

Communicating with a Social Robot

Whereas CASA supports the notion that people treat computers similarly to how they would treat other humans, interacting with a social robot presents a communication context unique to typical human-computer interaction (HCI) and significantly distinct from computer-mediated communication (CMC). To highlight this distinction, Zhao (2006) argues that CMC is the communication between humans through a computer medium, whereas HRI is the direct human communication with a computer (robot). The robot is not just a medium for communication; it is the intended receiver of communication (e.g., I speak *through* a mobile phone to my intended receiver, another human, but I speak face-to-face *with* a robot). The technology and its implications for communication differ drastically.

Additionally, HCI is more commonly used in reference to interaction with a computerized interface to optimize use of a larger technical system (Zhao, 2006). For example, HCI might look at how humans interact with and navigate through a web page. Ultimately, researchers might determine that because humans naturally categorize information, it makes sense to build the interface menu into categories for users to click through and navigate the system. Conversely, HRI incorporates more human-human communication characteristics that go beyond interacting with an interface. As Zhao states, “Humanoid social robots are not user-friendly computers that operate as machines; rather, they are user-friendly computers that operate as humans.” (2006, p. 403). Breazeal (2001) defines a social robot as an autonomous entity designed specifically to interact with humans. The addition of “humanoid” implies that these robots interact in a human-like manner, adhering to social rules and expectations. As many researchers would then argue, it is important that HRI studies be classified distinctly from HCI as the capabilities and implications of each field differ significantly.

This distinction does not, however, indicate that the CASA paradigm is irrelevant to the field of HRI. If anything, it may be assumed that the results predicted by CASA could be stronger, eliciting a more human-like interaction, toward a social robot, which provides more social cues than a simple computer. According to Kramer, von der Putten, and Eimler (2012), people engage in communicative behaviors in a manner consistent with how they typically communicate in human-human interactions as soon as an entity appears to be sufficiently social. The authors go further to suggest that humans are incapable of communicating differently than how their prior experiences have conditioned them to communicate. Instead, when people encounter an unknown entity such as a robot, they project existing social schemas onto them in order to know how to interact with them. Therefore, it is expected that humans will engage communicative acts with a robot in a similar manner to which they engage communicative acts with another human. This has been demonstrated in prior research linking the CASA paradigm to HRI (Lee, Park, & Song, 2005; Lee, Peng, Jin, & Yan, 2006; Park, Kim, & del Pobil, 2011)

Duffy (2003) contends that the current state of robots is largely a game of how well the robot can cheat, fooling humans into thinking that it is human-like when, in fact, its flaws begin to surface as individuals have more time to interact with it. Additionally, research has demonstrated that before interaction with a robot, people anticipating the interaction report greater uncertainty as well as expectations of lower social presence and social attraction compared to anticipating an interaction with a human (Spence, Westerman, Edwards, & Edwards, 2014). Regardless, along a spectrum of human-likeness, a social robot can be expected to exhibit more human-like characteristics than a computer, both in appearance and behavior. As the level of sophistication in robotics continues to grow as it has in the decade since Duffy's article, the level of anthropomorphism, or human-likeness, can be expected to increase as well. The degree of robotic presence in society will likely also influence perceptions and reduce levels of uncertainty felt in interacting with social robots.

Negotiating with Computerized Entities

In order to draw a connection between guilt and its role in human-robot negotiation, the following section highlights prior research related to HCI and HRI negotiation. Previous research has examined the difference in how humans interact with computerized entities, any non-human object that functions based on computer programming, in a negotiation situation. Torta and colleagues (2013) studied these differences by incorporating the 'Ultimatum Game' (UG), a simple conflict game commonly used in Game Theory research that places two entities as opponents in a scenario where a sum of goods or money is distributed between them. Before playing the game, participants were shown one of three randomly assigned photographs (human, humanoid robot, computer) and informed that this would be their opponent. Results indicated that humans tended to treat other humans and humanoid robots similarly, responding to distributive offers, offers intended to achieve concessions from the other party (Pruitt, 1981), more quickly and rejecting low offers more often compared to when responding to a computer opponent. This study suggests similarities between HHI (human photograph) and HRI (humanoid robot photograph) that can be elicited even through minor exposure to the opponent, which in this case, involved a simple photograph of the other party. This study demonstrates that participants make and maintain associations regarding who they think is negotiating on the other end of the computer.

Other studies related to HCI have examined the role of emotions expressed by a computer agent on how humans respond in a negotiation (de Melo, Carnevale, & Gratch, 2011, 2012). In these studies, researchers incorporated a negotiation game similar to one used in studies examining human-human negotiation (Van Kleef et al., 2006). Humans participate with a virtual agent in a complex negotiation game that incorporates negotiation of three issues related to the sale of a cellular phone: price, warranty, and service contract. Participants negotiate for multiple rounds, which allows for counter offers and feedback from the virtual agent in the form of verbal and nonverbal expressions. Results from these studies support claims that emotions play a similar role in HCI, and potentially HRI, as they do in HHI (de Melo et al., 2011, 2012). This stands to reason when considering Nass's CASA paradigm. Emotions are complex human characteristics, which when expressed by a computer or a robot, they provide social cues that trigger human-like attributions. The authors argue that people use emotions as a way to determine where the limits exist in negotiation (Van Kleef et al., 2006). For example, an expression of anger might indicate that the negotiations are nearing the opponent's limit, so an individual might begin to make concessions in order to keep the negotiation from ending without resolution. Guilt has been demonstrated to have an effect when portrayed by a virtual agent, such that participants concede more to a guilty virtual agent than to an expressionless agent (de Melo et al., 2012). Although de Melo and co-authors demonstrate emotions such as guilt can affect a human-nonhuman negotiation, guilt in the aforementioned study is being portrayed by the virtual agent, not elicited in the human opponent. The proposed study will flip these roles such that the robot is not expressing guilt, but rather seeking to inspire guilt in the participant opponent.

It remains unknown as to whether a robot can inspire such emotions in a human through emotional appeal. However, research has demonstrated that humans are highly effective at recognizing social cues (Nass, Steuer, & Tauber, 1994; de Melo et al., 2012) such that they can accurately identify human communication and emotions through non-human entities. Additionally, research has shown that humans will sympathize with robots and that this sympathy increases with more anthropomorphic robots (Riek, Rabinowitch, Chakrabarti, & Robinson, 2009). Knowing this and using Nass's CASA paradigm, which suggests that humans treat computers similarly to how they treat other humans, hypotheses may be formulated that incorporate human-human negotiation theory as predictors of human-robot negotiation behavior. Therefore, Face Negotiation Theory and the findings regarding guilt as an indicator of obliging and conceding behavior in U.S. participants (Zhang et al., 2014) are expected to transfer to a human-robot negotiation. As a result, the following hypothesis is offered.

H1: Humans will concede more to a robot negotiator communicating guilt appeals than to a robot communicating strictly logical appeals.

This hypothesis is supported by FNT, which suggests that in order to mitigate the risk of damaging one's face, one will engage in facework, employing an obliging conflict style, which involves making more concessions to counteract perceived transgressions of unfairness against the robot. Although some research has suggested that the CASA paradigm is limited and that empathy toward robots that have been treated unfairly or abused is not as strong as when it occurs to a human (Bartneck, Rosalia, Menges, & Deckers, 2005), a robot using emotional appeals should still be able to elicit greater feelings of guilt than a robot employing logical appeals. However, this raises the question as to what effect the robot's perceived agency in negotiation has on a human's behaviors in negotiation. Will someone concede more to a robot negotiating on behalf of another human (as an agent) or one negotiating on its own behalf (as a principal)? The following discussion of agency is based on terms of agent, a party in charge of negotiation and decision-making on behalf of another, and a principal, a party directly affected by the negotiation, borrowed from Spremann (1987).

As suggested by Bartneck and colleagues (2005), individuals may attribute more empathy and feel more guilt if they feel they have ultimately transgressed against another human. Therefore, if a robot is serving as a third party negotiator *for* a human, a participant may associate any personal decisions and tactics used as directly affecting the human, *not* the robot. Additionally, research conducted by Pruitt and Johnson (1970) revealed that parties tended to concede more to a third party negotiator (agent) than directly to the opposing party (principal), because they felt there was less need to save face. In other words, conceding directly to a principal is seen as more face threatening than conceding to an agent. Again, this highlights the notion that the robot negotiator is unlikely to be viewed as the direct opponent, but rather as a less face-threatening medium for negotiating with the true opponent. This has been previously supported in the context of human-human negotiation, finding that negotiations that incorporated the use of an agent resulted in better outcomes for the principal than if the principal alone was the sole negotiating party (Bazerman, Neale, Valley, Zajac, & Kim, 1992). Expectations may also play a significant role in how an individual negotiates with a robot agent (third-party negotiator). A partisan agent that favors the opposition tends to influence perceptions such that participants expect to achieve less in the negotiation (Conlon & Ross, 1993). This expectation to receive less may influence participants to be more willing to make concessions in negotiation.

Although the previous literature seems to suggest that a robot agent would receive more concession than a robot principal, other studies have found inconsistencies with how people cooperate with robots compared to humans. For example, Rilling, Gutman, Zeh, Pagnoni, Berns, and Kilts (2002) showed that people tended to cooperate more with humans than computers and that certain regions of the brain were only activated when playing a negotiation game against another human. Conversely, Kircher and colleagues (2009) found no significant differences in how participants cooperated with a computer compared to a human, despite similar brain activity findings. Further still, Sanfey, Rilling, Aronson, Nystrom, and Cohen (2003) found that participants accepted unfair offers more often from a computer than from another human. Collectively, the inconsistencies of this research and the potentially competing implications with the literature on agency vs. principal negotiation raises questions as to whether the agency effect of the robot will be able to overcome the effect of the embodied robot itself. As a result, two research questions are posed, one regarding the potential effect of robot agency and another addressing a potential interaction effect between the independent variables.

RQ1: What effect does robot agency have on an individual's level of concession in a negotiation scenario?

RQ2: What interaction effect exists between robot agency and robot appeal on participant concession?

CHAPTER III

METHODOLOGY

In order to test the proposed hypothesis and research questions, a 2(Principal, Agent) x2(Logic Appeal, Guilt Appeal) factorial experiment was conducted. Manipulating these four conditions in a simulated negotiation game sought to test whether H1) humans will concede more to a robot negotiator communicating guilt appeals than to a robot communicating no emotional appeals. Additionally, the experiment sought to answer two research questions. RQ1: What effect does robot agency (principal or agent) have on an individual's level of concession in a negotiation scenario? RQ2: What interaction effect exists between robot agency and robot appeal?

Participants

The sample consisted of 84 undergraduate students currently enrolled in either the introductory communication course (COM 1000) or the communication inquiry course (COM 2010) at Western Michigan University. Of the participants, 72.60% (n = 61) were female, while 27.40% (n = 23) were male. The majority (64.30%, n = 54) identified as Caucasian, followed by African-American (15.50%, n = 13), Other (8.30%, n = 7), Hispanic (6.00%, n = 5), Asian (4.80%, n = 4), and Native American (1.20%, n = 1). Participant ages ranged from 18 to 47 years, with a mean of 21.27 (SD = 4.60) and a median of 20 years. Participants were recruited via convenience sampling and invited to participate by the primary investigator through use of the WMU SONA research system. The SONA research system is an online tool used by student and faculty researchers to solicit anonymous student participation in campus-wide research studies. Upon participation, each participant was provided with SONA research credit, which were in turn applied to a required class research assignment. An alternative assignment was provided to those students choosing not to participate in studies through SONA.

Stimulus Material

ANeRo

During the experiment, participants interacted with a telepresence robot, named the Autonomous Negotiation Robot (ANeRo), produced by Double Robotics. ANeRo is remotely controlled via the Double Application on a computer or mobile device and moves via a self-balancing gyroscopic wheelbase, similar to the technology used in Segway transportation devices. The Double Application controls the robot through an iPad, which is docked in the “head” of the robot, serving as the primary mode of visual stimuli or “face” of the robot. The robot utilizes two cameras, both borrowed from the iPad in order for the operator to look forward and down. Additionally, the height at which ANeRo stands is remotely adjustable varying from 47” to 60”. In this experiment, the height was adjusted to its lowest point, 47”, so that it remained at eye level with participants, who were seated throughout the experiment. For the purposes of this experiment, the telepresence robot employed vocal messages, using a gender-neutral, computer-synthesized voice, as well as text-based messages that appeared across the screen of the robot when presenting offers and feedback during negotiation.

Wizard of Oz

It was important for the robot to appear as though it was functioning on its own, without the control of a human. It was named the Autonomous Negotiation Robot to reflect this ideal state of an autonomous robot with specific expertise in managing negotiation. However, the Double telepresence robot is not an autonomous robot, as the name suggests. Therefore, in order to test the hypothesis and research questions, a Wizard of Oz (WoZ; Kelley, 1984) technique was employed by the researcher. This is a common technique used by researchers to simulate a human-robot or human-computer interaction (Fraser & Gilbert, 1991; Green, Huttenrauch, & Eklundh, 2004, Steinfeld, Jenkins, & Scassellati, 2009) as it has been argued that robot/computer technology is not yet sophisticated enough to simulate the interactions desired by researchers. Using this technique, the researcher remotely manipulates the behaviors of the robot interacting with a human, creating the illusion that the robot is behaving as an autonomous entity.

In order to create the illusion of autonomous behavior, the experiment required two lab spaces. The first lab space was designed to mediate the participants' interactions with the lab assistant and the robot. In a separate lab space, the control room, the researcher manipulated the robot's behaviors and monitored the interaction. The researcher was able to observe the interaction and collect data by utilizing the robot's cameras and through use of the microphone that had been discreetly planted on the robot. Because the Double Robot is a telepresence robot by design, the monitor traditionally uses the camera of the other participant to display their face; however, for this experiment, the control room camera was rerouted to reflect the images of a separate monitor via which the researcher displayed text. Audio of the computerized voice also originated from the control room computer, cued to play at the display of each text message, and was channeled through the robot's speakers. All of this was controlled in a darkroom environment to reduce glare and to make the video feed look as realistic as possible, as though it was being created directly via the robot rather than remotely on a different platform. To further establish perceived realism and autonomy of the robot, the lab assistant used commands such as, “ANeRo, please approach the seller” and “ANeRo, return home”, to cue the robot to move in and out of negotiation position. Additionally, ANeRo began each interaction by introducing itself and performing a simple greeting ritual with each participant.

The Negotiation Game

The purpose of this research was to examine how feedback and perceived opponent agency influence offers and counter offers in a negotiation. A negotiation task was needed that allows for multiple rounds of offers and counter offers beyond a single offer and immediate acceptance or rejection as allotted by the Ultimatum Game. As a result, a more complex negotiation game was selected from prior research conducted by Van Kleef, De Dreu, and Manstead (2004). In the more complex negotiation game, participants negotiate three issues in a mobile phone sale: price, warranty period, and duration of the service contract, which provides the aforementioned affordances.

During the negotiation game, the participant was assigned the role of the seller. As the seller, there is significant room for profit as the device for sale cost only $30 to manufacture, which makes the cost of service and warranty replacement fairly cheap as well. This is a modification from the previous game incorporated by Van Kleef and fellow researchers, but provides the participant with adequate knowledge of the ramifications associated with the sale. In this case, the seller can never truly lose money for his/her business. As a result of this addition to the game, a modification to the participant’s payoff chart (see Table 1) is made to maintain consistency regarding the fact that there is no possibility of receiving negative or zero points at any level of negotiation.

Table 1

Participant Payoff Chart

	Price of Phones		Warranty Period		Service Contract	
Level	Price ($)	Payoff	Warranty	Payoff	Service	Payoff
1	150	500	1 month	150	1 month	300
2	145	450	2 months	135	2 months	270
3	140	400	3 months	120	3 months	240
4	135	350	4 months	105	4 months	210
5	130	300	5 months	90	5 months	180
6	125	250	6 months	75	6 months	150
7	120	200	7 months	60	7 months	120
8	115	150	8 months	45	8 months	90
9	110	100	9 months	30	9 months	60

In earlier studies (Van Kleef et al., 2004, 2006; de Melo et al., 2011, 2012), nine levels of negotiation were used for each of the three elements available for negotiation. The goal of the seller is to get as many points as possible, with level one providing the highest point values and level nine providing the lowest. The previous studies had set level nine to zero points, but payoffs have been modified in this experiment such that the least amount of points a participant may receive at level nine are 100, 30, and 60 for negotiating price, warranty period, and service contract respectively. Even in the worst case scenario for the participant, the participant would concede to all level nine offers, representing a payoff of 190 points, which remains well above merely breaking even at zero points. If the participant perceives no threat to receiving negative points or no points at all, he/she may be more open to guilt appeals expressed by the robot. Again, the reasoning behind this modification is to control for perceptual influences derived from potentially receiving a blatantly negative outcome of “zero points”.

The game begins with the buyer (robot) making an initial offer. Throughout the negotiation, the buyer follows a predetermined pattern of offers for every unique participant in order to control for variances caused by inconsistent offers. The offers are proposed as follows (for price – warranty – service): Round 1, 8-7-8; Round 2, 8-7-7; Round 3, 8-6-7; Round 4, 7-6-7; Round 5, 7-6-6; Round 6, 6-6-6. According to previous research (De Dreu & Van Lange, 1995), this pattern of offers was found to have face validity and represented a middle ground between cooperativeness and competitiveness. Regardless of whether opponents reach agreement by the end of round six, the researcher will interrupt the game, informing the participant that time has expired. Additionally, if at any point in the game the participant presented an offer that is more favorable to the robot than the robot’s next predetermined offer, the robot immediately accepted, ending the game.

Pre-Measures

In considering potential factors that may influence an individual’s performance behaviors in the negotiation game, three factors were identified as potential covariates. The first covariate of interest is one’s predisposition to experience guilt, which may influence how affected a participant might be by the robot’s guilt appeals. The second and third covariates are related as two dimensions of self-construal, interdependence and independence, discussed earlier. It seems to follow that one's individual focus versus one's focus on the well being of the collective may co-vary with concession in negotiation. The scales used to measure these covariates are discussed further below.

Guilt and Shame Proneness (GASP) Scale

In order to discern whether there is a mediating effect of one's predisposition to experience guilt on their willingness to concede in the negotiation, the Guilt and Shame Proneness (GASP) Scale (Cohen, Wolf, Panter, & Insko, 2011) was administered to participants before their participation in the negotiation game. The 16-item scale is comprised of four subscales (negative behavior evaluations (NBE), guilt repair tendencies, negative self-evaluations (NSE), shame withdrawal tendencies), the first of which was of primary interest, because it is namely associated with the feeling of guilt due to negative evaluations of one's behavior. The other three sub scales respectively relate to one's tendency to make amends for negative behaviors, one's level of shame resulting from a negative evaluation of self, and one's tendency to withdraw from situations in which he/she feels shame. In seeking to reduce participant mortality risk, and as these three subscales are not wholly relevant to the study, they were excluded from the study. While still a fairly recent measure, prior studies have reported Cronbach's alphas of .69 and .71, demonstrating acceptable to good reliability for the NBE subscale (Cohen, et al., 2011). In this study, a reliability coefficient of .63 ($M = 23.19$, $SD = 3.91$) was obtained, demonstrating acceptable reliability.

Self-Construal Measures

In addition to an individual's proneness to make negative behavior evaluations and experience guilt, it is equally important to understand how one aligns on the self-construal measures of independence and interdependence. As described above in the discussion of FNT, self-construals are the individualistic counterparts of a larger cultural dynamic of collectivism vs. individualism. In order to measure these self-construal factors and determine whether they mediate any concessionary effects, Singelis's two 12-item measures of independent and interdependent self-construals (1994) were used as a pre-measure. Participants respond to the measures by indicating level of agreement on a 7-point Likert-type scale regarding statements such as "Speaking up during a class is not a problem for me" (Independent Measure) and "Even when I strongly disagree with group members, I avoid an argument" (Interdependent Measure). Initial construction of these measures indicated reliability coefficients of .73 and .74 for the interdependent measure and .69 and .70 for the independent measure. In this study, reliability coefficients of .68 and .69 were obtained respectively. Further analysis of individual items in the interdependent measure reliability revealed that removing a single item from the 12-item scale elevated the scale's reliability from .68 (acceptable) to .71 (good reliability). The particular item in question, "I should take into account my parents' advice when making education/career plans" was removed from analysis.

Post-Game Measures

The primary objective of this study is to assess the level of concessions an individual is willing to make to a robot. As such, the primary form of measure will include recorded observations of counter offers made by the human participant. During analysis, a score derived from the total point concession made over the course of the game will represent the dependent variable. Observations were mediated via the camera and microphone built-in to ANeRo and recorded by the researcher who watched the video feed and manipulated ANeRo's behaviors remotely in another room. Additional post-game measures[1] questionnaires were administered at the conclusion of the negotiation game measuring varying perceptions of the robot and negotiation outcomes. As these measures were not related to the proposed hypothesis or research questions, they have been excluded from this paper's discussion.

Procedure

In the convenience sample, recruited individuals attended a scheduled lab session where upon arrival, they were greeted by a lab assistant and provided with an iPad containing the informed consent document. Participants were randomly assigned to one of four conditions (Principal Guilt, Principal Logic, Agent Guilt, Agent Logic) based upon all available combinations of the independent variables. Participants reviewed the informed consent and were instructed that clicking to the next window on the iPad indicated consent. Next, the lab assistant informed participants that they were going to play a game with the Autonomous Negotiation Robot, ANeRo, which was positioned across from the participant in the lab space.

Before playing the game, participants responded to the GASP questionnaire and read a brief synopsis of the game rules on the iPad, in which the game is described and ANeRo and its motives (Negotiating on behalf of itself or on behalf of another human) are introduced to the participant (Appendix A & B). Participants were informed that during the negotiation, ANeRo could hear and understand their verbal offers and that as long as they spoke clearly, the robot would understand offers in the form of levels (e.g., 1, 1, 1) or corresponding category amounts (e.g., $150, 1 month warranty, 1 month service contract). Participants referenced the payoff chart while playing and were told not to deviate from the available options. Once participants provide a counter-offer, ANeRo processes the offer and responded with a counter-offer displayed across the robot's screen. Additionally, a description of the participant's role as seller is provided along with rules of the game. The research informed the participant that his/her robot opponent also has a payoff chart to follow and that the goal of the robot was to obtain the lowest phone price, the highest warranty in months, and the highest service contract in months.

Following review of game rules and introduction to the robot, participants began the negotiation game. Participants are informed at the beginning of the game that they have been randomly assigned to either the buyer or seller position and that the game will continue until either an agreement is reached or time expires. Despite the instructions given to subjects, all participants were automatically assigned to the seller role. The justification for this assignment was to ensure as much as possible that the robot is in a position to elicit guilt from the participant. For example, it may be unlikely for a consumer to enter a store with the intention of buying a phone and ultimately feel guilty about getting too good of a deal. Instead, it may be more plausible to presume that a seller with significant control over the final price of the phone and with advanced knowledge about marginal gains generated by the sale may feel guilt as a result of being inflexible to consumer emotional appeals for a lower price.

At the conclusion of the negotiation game, the researcher instructed the participant to respond to a final questionnaire regarding self-reported guilt, perceived credibility of the robot in the negotiation, overall satisfaction with the negotiation process and outcomes, and standard participant demographics (biological sex, age, ethnicity, etc.). Lastly, participants were debriefed about the nature of the study and thanked for their participation.

Manipulating Emotional Appeals

In order to determine whether robots can elicit guilt through emotional appeals, sets of behaviors needed to be established that effectively conveyed to the human that the robot felt that the human was being unfair or had in some other way transgressed against the robot. In order to create messages for each appeal condition (Guilt, Logic), the researcher drafted multiple iterations of messaging. To ensure message validity, final messages were coordinated through rounds of feedback from a panel of researchers, lead by a scholar in the area of interpersonal communication, until messages were felt to appropriate reflect the intended condition (e.g., guilt appeal, “I feel like you are trying to take advantage of me”; logic appeal, “This offer does not make financial sense for me”). In addition to field scholars and as a further manipulation check, select graduate and undergraduate students were asked to informally evaluate whether the selected messages were either sufficiently guilt inducing or emotion-free. Messages also had to be adjusted slightly to reflect the agency condition of the robot (Principal, Agent). See Appendix C for a complete set of condition messages.

CHAPTER IV

DATA ANALYSIS AND DISCUSSION

Results

In order to determine whether a significant difference existed between the independent variables (robot agency, negotiation appeal) regarding their effect on a human's concession to a robot during a negotiation task, a 2-way (2 x 2) factorial ANOVA (analysis of variance) test was conducted. Levene's test for equality of variance was not significant for the dependent variable (total point concession; $F = .331$, $p = .803$). No significant main effects were found among condition appeal, $F(1, 80) = .25$, $p = .618$, or condition agency, $F(1, 80) = .53$, $p = .470$, related to change in the dependent variable. Nor was there a significant interaction effect between the two condition variables, $F(1, 80) = .002$, $p = .969$. Initial testing did not support H1 or provide directional answers to RQ1 or RQ2. Table 2 details the item means and standard deviations on the dependent variable for each of the four unique conditions.

Table 2

Means and Standard Deviations for Total Point Concession

Dependent Variable	Condition							
	Principal Logic		Principal Guilt		Agent Logic		Agent Guilt	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total Point Concession	239.25	173.67	220.00	150.85	263.75	168.15	247.27	161.01

In order to account and control for potential mediating effects (proneness to negative behavior evaluations (guilt), interdependence self-construal, independence self-construal), three 2-way (2 x 2) factorial ANCOVA (analysis of covariance) tests were conducted. The ANCOVAs yielded similar non-significant results among all covariates for main effects among the two condition factors (appeal, agency) as well as an interaction effect between the condition factors, (negative behavior evaluations, $F(1, 79) = .33$, $p = .569$, $F(1, 79) = .53$, $p = .470$, $F(1, 79) = .003$, $p = .960$; interdependence, $F(1, 79) = .01$, $p = .916$, $F(1, 79) = .62$, $p = .434$, $F(1, 79) = .04$, $p = .853$; independence, $F(1, 79) = .55$, $p = .459$, $F(1, 79) = .19$, $p = .666$, $F(1, 79) = .02$, $p = .895$, respectively).

Discussion

The purpose of this research experiment was to determine whether robots could effectively manipulate emotions in a negotiation scenario such that they might influence a human's behavior. To test this idea, guilt appeals were used in contrast to logical appeals to determine if a significant difference in overall concession existed between the two variable attributes. Additionally, a second variable, robot agency, was introduced to examine what effect the robot's role (as principal negotiator or agent negotiating on behalf of a human) may have on willingness to concede. Ultimately, results indicated no significant differences among the independent variables. Rather, results showed that participants appeared to negotiate drastically different from one another based on some other unaccounted variable. This dramatic difference in negotiation style becomes evident in reviewing the item standard deviations on Table 2. Further discussion of results related to each independent variable is continued below, followed by a discussion of current study limitations and proposed direction for future research.

Guilt Eliciting Robots

Given that the results indicate no difference between a robot using guilt appeals versus one employing logical appeals, this raises interesting implications for the future role of robots in surrogate human positions. Although surprising, these results may speak to the relatively low familiarity most individuals have with human-robot interaction (Mutlu, Shiwa, Kanda, Ishiguro, & Hagita, 2009; Kulic & Croft, 2007). Most individuals may have a referent for "robot" stored in their perceptual set, but the reference may yet be underdeveloped such that values, attitudes, and beliefs surrounding these entities remain either unknown or highly indifferent. It may be possible to integrate robots into roles that eventually elicit emotional responses of empathy from humans, but it is perhaps too early to determine its effectiveness until the presence of robots become more substantial and commonplace. For example, several participants that interacted with the lab assistant at the end of the study posed questions such as, "Is this a new technology?", "How does it know what to do?", "How does it see me?" This demonstrates that participants clearly had uncertainty regarding the robot and were perhaps more thrown off by the unfamiliar and unique nature of the interaction than by any other factor.

The Computers as Social Actors (CASA) paradigm suggests that humans will treat computers, and similarly robots, in a manner similar to how they treat other humans (Nass, Steuer, Tauber, & Reeder, 1993). By this thinking and in conjunction with findings extended from Face Negotiation Theory (Ting-Toomey, 1988) that demonstrate guilt elicits an obliging conflict style (Zhang, Ting-Toomey, & Oetzel, 2014), one would expect guilt appeals used by a robot to inspire greater concession from a human negotiator. However, as briefly discussed in the literature review section, reciprocity is another important component of interpersonal and human-computer interaction (Katagiri, Nass, & Takeuchi, 2001; Nass, Moon, & Carney, 1999). Individuals will tend to reciprocate behaviors enacted by the computer (robot) in a back and forth, tit for tat, manner. Within the game, the offers provided by the robot were controlled in all conditions such that regardless of condition, the robot would always provide the same offer on the same corresponding turn. Although this pattern of offers had previously been determined to demonstrate neither an overly competitive nor an overly cooperative pattern, the offers change fairly systematically with a minor, one-level concession between each turn. It is possible that actual offers superseded the importance of the robot's message, overall, causing individuals to reciprocate a minor concession with a minor concession.

In addition to situational unfamiliarity, a lack of empathy for the robot may have played a significant role. In the post-test questionnaire, a 7-point, single-item question related to empathy was posed to participants, "How sorry do you feel for the robot negotiator?" A response of "1" indicated "Not sorry at all" and a response of "7" indicated "Extremely sorry". Means collected from the data set indicated that regardless of the variable attribute (Guilt, Logic, Principal, Agent), empathy barely exceeded a "2"; see Table 3. Previous research has indicated that empathy is correlated to guilt such that the greater empathy an individual feels toward another entity, the more prone they will be toward feeling guilt (Tangney, 1991). Hoffman (1984) further suggests that empathy foregrounds guilt, serving as the basis from which guilt develops. Therefore, if empathy is lacking, guilt, a face-threatening emotion (Zhang, Ting-Toomey, & Oetzel, 2014), does not manifest or require the enactment of face-saving strategies such as obliging or concession. It stands to reason then that if a person reports no empathy toward a robot, a guilt appeal from that robot is unlikely to have significant influence on negotiation results.

Empathy and assigning human attributions to the robot may also be inextricably linked to the three covariates (proneness to negative behavior evaluations, independence, and interdependence) examined in this study. Regardless of one's position along these three factors, all of which demonstrated acceptable to good reliability, the factors are related to one's relationship or interaction with another human. If participants did not feel the same empathy or similarity with the robot as with another human, it is unlikely that the covariates would have mediated any effect on concessions with a robot.

Table 3

Means and Standard Deviations for Empathy

Dependent Variable	Independent Variable Attribute							
	Guilt		Logic		Principal		Agent	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Empathy	2.07	1.37	1.38	.74	1.74	1.11	1.74	1.23

Without the apparent sense of empathy for the robot, participants instead expressed frustration with the robot. In the guilt appeal conditions, many of the participants made comments demonstrating this frustration that included, “If [ANeRo] was real, we would have been going at it”, “Stubborn little thing”, “[ANeRo] is a hard negotiator." "[ANeRo] is sassy", "I think [ANeRo] hates me", and "Sorry, Robot, for being mean. I was just getting annoyed." Similar comments, such as "This is a sassy little thing" even manifested in the logical appeal conditions, demonstrating that regardless of intent, participants may have been assigning a "difficult" personality to the robot based on its role as an opponent. This lack of empathy may explain why the current study did not indicate findings similar to Zhang and colleagues' study (2014) in which guilt elicited greater obliging and concession. However, this lack of empathy resulting in expressions of frustration may be a product of the messaging employed by the robot, which will be addressed further in the discussion on limitations.

Potential study manipulation issues cannot be entirely ruled out for influencing the results. However, it is likely that the non-significant findings accurately represent human perceptions of robots and the degree to which humans find robots deserving of empathy. Nass and Moon (2000) suggested that the attributions humans make to computerized entities are a result of mindless attribution. In order for this mindless attribution to occur, the interaction's cognitive demand cannot exceed a certain threshold. In other words, mindless attributions are dependent upon simple, heuristic cues that allow individuals to process the communication peripherally (non-critically) rather than centrally. Petty and Cacioppo provide an in-depth discussion of the central and peripheral processing of information within their development of the Elaboration Likelihood Model (ELM; 1986).

Neither guilt, an intense form of emotional distress (Baumeister, Stillwell, & Heatherton, 1994), nor a complex negotiation game allow for peripheral message processing or mindless attribution. Instead, scenarios such as the one used in this study require significant cognitive energy, which causes individuals to critically evaluate the messages and corresponding messenger. As a result, humans are forced to ask themselves, “Why should I feel guilty toward a robot?”. In actively processing this question, an awareness of the robot’s un-human and inorganic qualities surface, resulting in an acknowledgement that the robot is not deserving of human-like attributions such as empathy. This acknowledgement reflects responses Nass and Moon (2000) collected from participants asked to reflect on human-computer interaction. A significant majority emphatically expressed that computers are neither human nor deserving of human-like attributions and communication.

Robot Agency

In addition to the message appeal employed by a robot, this study sought to understand what effect the agency of a robot would have on overall human concession. Results from the current study indicate that there was no effect on concession whether the robot served as the principal negotiator or as an agent acting on behalf of a human. This finding further affirms the complexity of this issue. Although some literature would seem to support the notion that participants are more likely to concede to a robot as principal due to the third-party’s less face-threatening position (Pruitt and Johnson, 1970) and the participants lower expectations for success in dealing with a partisan third-party (Conlon & Ross, 1993), other research frames a different prediction. This other research seems to suggest that individuals would concede more to a robot detached from a human element based upon patterns of interactions that found people are more likely to accept unfair offers from a computer compared to another human (Sanfey, et al., 2003). Given the complexity of the issue, it may be that multiple elements are confounding with one another to dilute the effects or cancel one another out altogether. The most effective means to look at this phenomenon may be as an isolated variable in future studies in order to more clearly parse out and control for additional factors.

Limitations and Future Direction

In consideration of future research, this study highlighted some potential deficiencies in the experiment that may have limited the observed effects on participant concession with a robot. First, a limitation briefly discussed earlier is related to the messaging used in the guilt-inducing condition. When creating the messaging, a panel of scholars was convened to draft and vet messaging that would be effective in eliciting guilt. Although these messages appeared valid and effective at initial face value, they did not appropriately elicit the intended guilt based on the empathy check employed post-test. Although the manipulation check provided by students and an expert scholar appeared to generate sufficient messaging, further pilot testing may have provided an additional, beneficial manipulation check. It is possible that in practice, the messages came across as too heavy-handed, inspiring anger and frustration, which may have been an issue in the current study.

Additionally, the range of offers varied drastically from one participant to another, even within identical conditions. A larger sample would be advantageous to see whether a larger data set would provide a more normative distribution of offers. Lastly, lab experiments often face the issue of ecological validity, but the current study may have been increasingly susceptible to such issues. Interacting with a robot is already an unfamiliar, perhaps even surreal, experience for many. That unfamiliarity paired with a "game" scenario did not provide a real-life environment. This study may have benefitted from putting the robot negotiator in a public space where individuals could negotiate over something more tangible than a fake cell phone. As well, using an android robot, capable of enacting more human-like social cues than a telepresence robot, may elicit different responses.

Despite a lack of significant results and the limitations reviewed above, this study creates an excellent jumping off point for future research regarding human-robot interaction, and more specifically, human-robot negotiation. Before this study, little research existed related to robots interacting with humans in an oppositional role. The field of communication and HRI could benefit greatly by addressing some of the limitations mentioned above and by assessing what contexts are best suited for an adversarial robot.

One point for future study includes the examination of facework in human-robot interactions. Although empathy rated low throughout this experiment, are humans still prone to enact expressive rituals, attempting to protect the face of a robot teammate or opponent? Additional insight may be garnered as to what effect a robot's use of face-saving strategies, both self and other centered, has on a human's perception of the robot's competence, homophily, and familiarity.

Of particular interest would be finding a way to break through the ecological validity barrier that currently limits HRI research. Until robots are incorporated vastly in day-to-day proceedings, the genuineness of results may be brought into question. Admittedly, this may partially be left up to a waiting game as manufacturers and consumers alike begin to adopt these technologies on a greater scale. Meanwhile, there are other areas of study focused on physical design and representation of robots such as size, mobility, linguistic characteristics, and other factors that likely effect human communication with a robot. Addressing these design cues from a human perception and communication standpoint will help pave the way for designing more effective and more well liked robots.

Conclusion

Ultimately, the widespread adoption of robotics in our society is inevitable. Between the noted benefits of robots in healthcare, education, and other meaningful human environments, the possibilities for robotics are seemingly endless at this time. The robots available today are growing in sophistication, and consumers are already beginning to see task-oriented and social robots becoming available at affordable price points. As manufacturers and consumers begin to adopt these robot technologies more broadly in public, the workplace, and in the home, it will be interesting to see whether individuals embrace robots more for either their functional or social utility. Understanding how to raise the communicative effectiveness of these robots through their ability to read, understand, interpret, and reproduce social cues such as verbal and nonverbal emotions may be the difference between a robot that proliferates the market and one that sits idly on the shelf. It is no longer a game of whether or not the robot can do what you ask, but rather it is a game of whether or not the robot can realistically simulate the creation of meaning with you.

Endnotes

post-game measures[1] – Post-measures administered for separate analysis included a perceived credibility scale (McCroskey & Teven, 1999), a scale of interpersonal attraction (McCroskey & McCain, 1974), a modified scale of negotiation satisfaction (Graham, 1985), and a set of demographic questions.

REFERENCES

Allred, K. G., Mallozzi, J., Matsui, F., & Raia, C. P. (1997). The Influence of Anger and Compassion on Negotiation Performance. *Organizational Behavior and Human Decision Processes, 70*, 175-187.

Bartneck, C., Chioke, R., Menges, R., & Deckers, I. (2005). Robot Abuse – A Limitation of the Media Equation. *Proceedings of the Interact 2005 Workshop on Abuse*, Rome.

Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. *American Economic Review, Papers and Proceedings, 97*, 170-76. doi:10.1257/aer.97.2.170

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*(2), 243-267. doi:10.1037/0033-2909.115.2.243

Bazerman, M. H., & Neale, M. A. (1983). Heuristics in negotiation: Limitations to effective dispute resolution. In M. H. Bazerman & R. J. Lewicki (Eds.), *Negotiating in Organizations.* 51-67. Beverly Hills, CA: Sage.

Bazerman, M. H., Neale, M. A., Valley, K. L., Zajac, E. J., & Kim, Y. M. (1992). The effect of agents and mediators on negotiation outcomes. *Organizational Behavior and Human Decision Processes, 53*, 55-73. doi:10.1016/0749-5978(92)90054-B

Bodtker, A. M., & Jameson, J. K. (2001). Emotion in conflict formation and its transformation: Application to organizational conflict management. *The International Journal of Conflict Management, 12*, 259-275. doi:10.1108/eb022858

Breazeal, C.L. (2001) *Designing Sociable Robots*. Cambridge, MA: MIT Press.

Breazeal, C. (2011). Social robots for health applications. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE.* Piscataway, NJ, USA, pp. 5368-5371.

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. *Proceedings of IROS*, Barcelona, Spain.

Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive robots in elderly care: A review. *Gerontechnology, 8*(2), 94-103. doi:10.4017/gt.2009.08.02.002.00

Broekens, J., Jonker, C. M., & Meyer, J.-J., C. (2010). Affective negotiation support systems. *Journal of Ambient Intelligence and Smart Environments, 2*(2), 121-144. doi:10.3233/AIS-2010-0065

Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology, 100*(5), 947-966. doi:10.1037/a0022641

Conlon, D. E., & Ross, W. H. (1993). The effects of partisan third parties on negotiator behavior and outcome perceptions. *Journal of Applied Psychology, 78*(2), 280-290. doi:10.1037/0021-9010.78.2.280

Covert, M. V., Tangney, J. P., Maddux, J. E., & Heleno, N. M. (2003). Shame-proneness, guilt-proneness, and interpersonal problem solving: A social-cognitive analysis. *Journal of Social and Clinical Psychology, 22*, 1–12. doi:10.1521/jscp.22.1.1.22765

De Dreu, C. K. W., & Van Lange, P. A. M. (1995). The impact of social value orientations on negotiator cognition and behavior. *Personality and Social*

Psychology Bulletin, 21(11), 1178-1188.

de Melo, C. M., Carnevale, P., & Gratch, J. (2011). The effect of expression of anger and happiness in computer agents on negotiations with humans. *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'11),* Taipei, Taiwan. Retrieved from http://ict.usc.edu/pubs/The%20Effect%20of%20Expression%20of%20Anger%20and%20Happiness%20in%20Computer%20Agents%20on%20Negotiations%20with%20Humans.pdf

de Melo, C. M., Carnevale, P., Gratch, J. (2012). The effect of virtual agents' emotion displays and appraisals on people's decision making in negotiation. In: Nakano, Y., Neff, M., Paiva, A., Walker, M. (eds.) IVA 2012. LNCS (LNAI), vol. 7502, pp. 53–66. Springer, Heidelberg

Deutsch, M. (1973). *The Resolution of Conflict*. New Haven CT, Yale University Press.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*, 177-190. doi:10.1016/S0921-8890(02)00374-3

Dufwenberg, M. (2002). Marital investments, time consistency, and emotions. *Journal of Economic Behavior & Organization, 48*, 57-69. doi:10.1016/S0167-2681(01)00225-6

Edwards, C., Spence, P. R., Gentile, C. J., Edwards, A., & Edwards, A. (2013). How much Klout do you have...A test of system generated cues on source credibility. *Computers in Human Behavior, 29*(5), A12-A16. doi:10.1016/j.chb.2012.12.034

Fogg, B. J., & Nass, C. (1997). Silicon sycophants: the effects of computers that flatter. *International Journal of Human-Computer Studies, 46*, 551-561. doi:10.1006/ijhc.1996.0104

Fraser, N. M., & Gilbert, G. N. (1991). Simulating speech systems. *Computer Speech & Language*, *5*(1), 81–99. doi:10.1016/0885-2308(91)90019-M

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication* (RO-MAN 2003). Oct. 31-Nov. 2, Milbrae, CA, 55- 60.

Graham, J. L. (1985). Cross-cultural marketing negotiations: A laboratory experiment. *Marketing Science*, *4*(2), 130-146. doi:10.1287/mksc.4.2.130

Graham, J. L., Kim, D. K., Lin, C.-Y., & Robinson, M. (1988). Buyer-seller negotiations around the pacific rim: differences in fundamental exchange processes. *The Journal of Consumer Research*, *15*(1), 48-54. doi:10.1086/209144

Green, A., Huttenrauch, H., & Eklundh, K. S. (2004). Applying the Wizard-of-Oz framework to cooperative service discovery and configuration. *In 13th IEEE International Symposium on Robot and Human Interactive Communication (Ro-MAN)* (pp. 575–580). IEEE. doi:10.1109/ROMAN.2004.1374824

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*, 151-181. doi:10.1207/s15327051hci1901&2_7

Hoffman, M. L. (1982). Development of prosocial motivation: Empathy and guilt. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 281-313). San

Diego, CA: Academic Press.

Hutson, S., Lim, S. L., Bentley, P. J., Bianchi-Berthouze, N., & Bowling, Ann. (2011). Investigating the suitability of social robots for the wellbeing of the elderly. *Proceedings of ACII 2011*, pp. 578-587. doi:10.1007/978-3-642-24600-5_61

Kahn, P. H., Freier, N. G., Friedman, B., Severson, R. L., & Feldman, E. N. (2004) Social and moral relationships with robotic others. *Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication*, Kurashiki, Okayama, Japan, September 20-22, pp. 545-550.

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), 61-84. doi:10.1207/s15327051hci1901&2_4

Kanda, T., Ishiguro, H. (2005). Communication Robots for Elementary Schools. In: Proceedings of the Symposium on Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (AISB 2005), pp. 54–63. Hatfield, UK.

Katagiri, Y., Nass, C., & Takeuchi, Y. (2001). Cross-cultural studies of the computers are social actors paradigm: The case of reciprocity. *In Proceedings of the Ninth International Conference on Human-Computer Interaction 2001*. 1558-1562.

Kaupp, T., Makarenko, A., & Durrant-Whyte, H. (2010). Human-robot communication for collaborative decision making – A probabilistic approach. *Robotics and Autonomous Systems*, *58*(5), 444-456. doi:10.1016/j.robot.2010.02.003

Kelley, J. F. (1984). An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems, 2*(1), 26–41. doi:10.1145/357417.357420

Ketelaar, T., & Au, W. T. (2003). The effects of feelings of guilt on the behavior of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion, 17*(3), 429-453. doi:10.1080/02699930143000662

Kircher, T., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., van Os, J. & Krach, S. (2009). Online mentalising investigated with functional MRI. *Neuroscience letters, 454*(3), 176-181.

Kramer, N. C., von der Putten, A., & Eimler, S. (2012). Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In M. Zacarias, & J. V. de Oliveira (Eds.) *Human-Computer Interaction: The Agency Perspective.* Springer Berlin Heidelberg.

Kulic, D., & Croft, E. A. (2007). Affective state estimation for human-robot interaction. *IEEE Transactions on Robots, 23*(5), 991-1000. doi:10.1109/TRO.2007.904899

Lee, E.-J., Nass, C., & Brave, S. (2000). Can computer-generated speech have gender? An experimental test of gender stereotypes. *Paper presented at the CHI 2000,* The Hague, The Netherlands.

Lehman, H., Iacono, I., Dautenhaun, K., Marti, P., & Robins, B. (2014). Robot companions for children with down syndrome: A case study. *Interaction Studies, 15*(1), 99-112. doi:10.1075/is.15.1.04leh

Leith, K. P., & Baumeister, R. F. (1998). Empathy, shame, guilt, and narratives of interpersonal conflict: Guilt-prone people are better at perspective taking. *Journal of Personality, 66*(1), 1-37.

Mandel, D., & Dhami, M. (2005). ‘‘What I did’’ versus ‘‘what I might have done’’:

Effect of factual versus counterfactual thinking on blame, guilt, and shame in prisoners. *Journal of Experimental Social Psychology, 41*, 627-35. doi:10.1016/j.jesp.2004.08

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224–253. doi:10.1037/0033-295X.98.2.224

McCroskey, J. C. & McCain, T. A. (1974). The measurement of interpersonal attraction. *Speech Monographs, 41*, 261-266.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs*, *66*, 90-104.

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human-robot conversations: How robots might shape participant roles using gaze cues. *In Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction.* ACM. doi:10.1145/1514095.1514109

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies, 45*, 669-678. doi:10.1006/ijh.01996.0073

Nass, C., & Lee, K., M. (2000). Does computer-generated speech manifest personality? An experimental test of similarity-attraction. *Proceedings of CHI 2000,* The Hague, Amsterdam.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*, 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., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, *43*, 223-239.

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

Nass, C., Reeves, B., & Leshner, G. (1996). Technology and roles: A tale of two TVs. *Journal of Communication*, *46*(2), 121-128.

Nass, C. I., Steuer, J. S., Tauber, E. & Reeder, H. (1993). Anthropomorphism, agency, and ethopoeia: Computers as social actors. *Proceedings of the International CHI Conference*. Amsterdam, The Netherlands. doi:10.1145/259964.260137

Oetzel, J. G. (1998). The effects of self-construals and ethnicity on self-reported conflict styles. *Communication Reports*, *11*(2), 133-144. doi:10.1080/08934219809367695

Oetzel, J., Meares, M., Myers, K. K., & Lara, E. (2003). Interpersonal conflict in organizations: Explaining conflict styles via face negotiation theory. *Communication Research Reports*, *20*(2), 106-115. doi:10.1080/08824090309388806

Oetzel, J. G., & Ting-Toomey, S. (2003). Face concerns and facework during conflict: A test of the face-negotiation theory. *Communication Research*, *30*, 599–624.

doi:10.1177/0093650203257841.

Oetzel, J., Ting-Toomey, S., Masumoto, T., Yokochi, Y., Pan, X., Takai, J., & Wilcox, R. (2001). Face and facework in conflict: A cross-cultural comparison of China, Germany, Japan, and the United States. *Communication Monographs, 68*(3), 235-258. doi:10.1080/03637750128061

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-205.

Pruitt, D. G., & Carnevale, P. J. (1993). *Negotiation in social conflict.* Pacific Grove, CA: Brooks/Cole.

Pruitt, D. G., & Johnson, D. F. (1970). Mediation as an aid to face saving in negotiation. *Journal of Personality and Social Psychology, 14*(3), 239-246.

Rahim, M. A. (1983). A measure of styles of handling interpersonal conflict. *Academy of Management Journal, 26,* 368–376.

Riek, L., Rabinowitch, T., Bremner, P., Pipe, A. G., Fraser, M. and Robinson, P. (2010) Cooperative gestures: Effective signaling for humanoid robots. *5th ACM/IEEE International Conference on Human-Robot Interaction (HRI),* 2010 . pp. 61-68.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social comparison. *Neuron, 35*(2), 395-405.

Robins, B., Dautenhahn, K., Boekhorst, R., & Billard, A. (2005). Robotic assistants in therapy and education of children with autism: Can a small humanoid robot help encourage social interaction skills? *Universal Access in the Information Society, 4*(2), 105-120. doi:10.1007/s10209-005-0116-3

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science, 300*(5626), 1755-1758.

Scassellati, B. (2007). How social robots will help us to diagnose, treat, and understand autism. *Robotics Research,* 28, 552-563.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*(5), 580-591. doi:10.1177/0146167294205014

Spence, P. R., Westerman, D., Edwards, C., & Edwards, A. (2014). Welcoming our robot overlords: Initial expectations about interaction with a robot. *Communication Research Reports, 31*(3), 272-280. doi:10.1080/08824096.2014.924337

Spremann, K. (1987). Agent and principal. In G. Bamberg & K. Spremann (Eds.), *Agency Theory, Information, and Incentives* (3-38). Berlin: Springer Verlag.

Steinfeld, A., Jenkins, O. C., & Scassellati, B. (2009). The Oz of Wizard: Simulating the human for interaction research. *In Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction.* ACM. doi:10.1145/1514095.1514115

Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences, 104*(46), 17954-17958. doi:10.1073/pnas.0707769104

Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology, 61*, 598-607. doi:10.1037/0022-3514.61

Thomas, K.W., & Kilmann, R. H. (1974).*Thomas-Kilmann conflict MODE instrument.* Tuxedo, NY: Xicom.

Ting-Toomey, S. (1988). Intercultural conflict styles: A face-negotiation theory. In Y. Y. Kim &W. Gudykunst (Eds.),*Theories in intercultural communication* (pp. 213–235). Newbury Park, CA:Wadsworth.

Ting-Toomey, S. (2005).The matrix of face: An updated face-negotiation theory. In W. B. Gudykunst (Ed.), *Theorizing about intercultural communication* (pp. 71–92).Thousand Oaks, CA: Sage.

Ting-Toomey, S., Gao, G., Trubisky, P., Yang, Z., Kim, H. S., Lin, S., & Nishida, T. (1991). Culture, face maintenance, and styles of handling interpersonal conflict: A study in five cultures. *International Journal of Conflict Management, 2*, 275-296. doi:10.1108/ eb022702

Ting-Toomey, S., & Kurogi, A. (1998). Facework competence in intercultural conflict: An updated face-negotiation theory. *International Journal of Intercultural Relations*, *22*(2), 187-225.

Ting-Toomey, S., Oetzel, J., & Yee-Jung, K. (2001). Self-construal types and conflict management styles. *Communication Reports, 14*, 87–104. doi:10.1080/08934210109367741

Ting-Toomey, S., Yee-Jung, K. K., Shapiro, R. B., Garcia, W., Wright, T. J., & Oetzel, J. G. (2000). Ethnic/cultural identity salience and conflict styles in four US ethnic groups. *International Journal of Intercultural Relations, 24*, 47-81. doi:10.1016/S0147-1767(99)00023-1

Torta, E., van Dijk, E., Ruijten, P. A. M., & Cuijpers, R. H. (2013). The ultimatum game as measurement tool for anthropomorphism in human-robot interaction. *Social Robotics*, 8239, 209-217. doi:10.1007/978-3-319-02675-6_21

Trubinsky, P., Ting-Toomey, S., & Lin, S.-L. (1991). The influence of individualism-collectivism and self-monitoring on conflict styles. *International Journal of Conflict Relations, 15*, 65-84. doi:10.1016/0147-1767(91)90074-Q

Tung, F.-W., Sato, K., Deng, Y.-S., & Lin, D.-Y. (2009). A cross-cultural study on the perception of sociability within human-computer interaction. *Proceedings of the HCI International Conference 2009*, San Diego, CA, USA, pp. 135-144. doi:10.1007/978-3-642-02767-3_15

Van Kleef, G. A., De Dreu, C. K.W., & Manstead, A. S. R. (2004). The interpersonal effects of emotions in negotiations: A motivated information processing approach. *Journal of Personality and Social Psychology, 87*(4), 510-528. doi:10.1037/0022-3514.87.4.510

Van Kleef, G. A., De Dreu, C. K.W., & Manstead, A. S. R. (2006). Supplication and appeasement in conflict and negotiation: The interpersonal effects of disappointment, worry, guilt, and regret. *Journal of Personality and Social Psychology, 91*, 124–142. doi:10.1037/0022-3514.91.1.124.

Westerman, D., Spence, P. R., & Van Der Heide, B. (2012). A social network as information: The effect of system generated reports of connectedness on credibility on Twitter. *Computers in Human Behavior*, *28*(1), 199-206. doi:10.1016/j.chb.2011.09.001

Zhang, Q., Ting-Toomey, S., & Oetzel, J. G. (2014). Linking emotion to the conflict negotiation theory: A U.S.-China investigation of the mediating effects of anger, compassion, and guilt in interpersonal conflict. *Human Communication Research, 40*, 373-395. doi:10.1111/hcre.12029

Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Society and Media, 8*(3), 401-419. doi:10.1177/1461444806061951

Appendix A

Robot Principal Game Description

Appendix A

Robot Principal Game Description

Today, you will be playing a negotiation game with an Autonomous Negotiation Robot, ANeRo, who will be serving as your opponent. The objective of the game you will be playing with ANeRo is to negotiate the sale of a new smart phone, the MAX 3S. **You will be negotiating 3 aspects of the sale: price, the duration of the warranty period, and the duration of the service contract**. For this scenario, you have been randomly assigned the role of "seller", and ANeRo has been assigned the role of "buyer". As the seller, your objective is to sell the phone at the highest price, with the lowest duration of warranty period and service contract. The cost to manufacture the MAX 3S is only $30, which as a result means replacing the phone and performing service is relatively inexpensive for you as the seller. As a result, even at its lowest price point, you, the seller, receive significant profit for every phone you sell.

The individual price points and durations of warranty and service are broken into levels as shown in the payoff chart below. As you can see from the payoff chart, as the seller, **your best-case scenario is to negotiate the highest price, the shortest warranty period, and the shortest service contract**, because it offers the greatest point payoff (500-150-300) for a total of 950 points. However, you may make any combination of offer you like in order to reach an agreement. You will be alternating turns with ANeRo to reach an agreed upon sale of the phone. ANeRo also has a programmed payoff chart, but it is different than yours. ANeRo will be negotiating for the lowest price, the longest warranty period, and the longest service contract. The game will continue either until you and ANeRo reach an agreement or until time expires. The game will begin with the buyer, ANeRo, making an initial spoken offer, which will also appear as text across the robot's screen. When making a counter-offer, speak clearly to ANeRo. The robot's voice recognition software is advanced, but it may still ask you to repeat your offer if it does not understand. You have been provided with a blank piece of paper where you may track yours and the robot's offers.

Payoff Chart

	Price of Phone		Warranty Period		Service Contract	
Level	Price	Payoff	Warranty	Payoff	Service	Payoff
1	$150	**500pts**	1 month	**150pts**	1 month	**300pts**
2	$145	**450pts**	2 months	**135pts**	2 months	**270pts**
3	$140	**400pts**	3 months	**120pts**	3 months	**240pts**
4	$135	**350pts**	4 months	**105pts**	4 months	**210pts**
5	$130	**300pts**	5 months	**90pts**	5 months	**180pts**
6	$125	**250pts**	6 months	**75pts**	6 months	**150pts**
7	$120	**200pts**	7 months	**60pts**	7 months	**120pts**
8	$115	**150pts**	8 months	**45pts**	8 months	**90pts**
9	$110	**100pts**	9 months	**30pts**	9 months	**60pts**

Appendix B

Robot Agent Game Description

Appendix B

Robot Agent Game Description

Today, you will be playing a negotiation game with an Autonomous Negotiation Robot, ANeRo, who will be serving as your opponent and will be negotiating on behalf of a human client. The objective of the game you will be playing with ANeRo is to negotiate the sale of a new smart phone, the MAX 3S. **You will be negotiating 3 aspects of the sale: price, the duration of the warranty period, and the duration of the service contract**. For this scenario, you have been randomly assigned the role of "seller", and ANeRo has been assigned the role of "buyer". As the seller, your objective is to sell the phone at the highest price, with the lowest duration of warranty period and service contract. The cost to manufacture the MAX 3S is only $30, which as a result means replacing the phone and performing service is relatively inexpensive for you as the seller. As a result, even at its lowest price point, you, the seller, receive significant profit for every phone you sell.

The individual price points and durations of warranty and service are broken into levels as shown in the payoff chart below. As you can see from the payoff chart, as the seller, **your best-case scenario is to negotiate the highest price, the shortest warranty period, and the shortest service contract**, because it offers the greatest point payoff (500-150-300) for a total of 950 points. However, you may make any combination of offer you like in order to reach an agreement. You will be alternating turns with ANeRo to reach an agreed upon sale of the phone. ANeRo also has a programmed payoff chart, but it is different than yours. ANeRo will be negotiating for the lowest price, the longest warranty period, and the longest service contract. The game will continue either until you and ANeRo reach an agreement or until time expires. The game will begin with the buyer, ANeRo, making an initial spoken offer, which will also appear as text across the robot's screen. When making a counter-offer, speak clearly to ANeRo. The robot's voice recognition software is advanced, but it may still ask you to repeat your offer if it does not understand. You have been provided with a blank piece of paper where you may track yours and the robot's offers.

Payoff Chart

	Price of Phone		Warranty Period		Service Contract	
Level	Price	Payoff	Warranty	Payoff	Service	Payoff
1	$150	**500pts**	1 month	**150pts**	1 month	**300pts**
2	$145	**450pts**	2 months	**135pts**	2 months	**270pts**
3	$140	**400pts**	3 months	**120pts**	3 months	**240pts**
4	$135	**350pts**	4 months	**105pts**	4 months	**210pts**
5	$130	**300pts**	5 months	**90pts**	5 months	**180pts**
6	$125	**250pts**	6 months	**75pts**	6 months	**150pts**
7	$120	**200pts**	7 months	**60pts**	7 months	**120pts**
8	$115	**150pts**	8 months	**45pts**	8 months	**90pts**
9	$110	**100pts**	9 months	**30pts**	9 months	**60pts**

Appendix C

Robot Condition Messages

Appendix C

Robot Condition Messages

Logical/Self-Agency	**Emotional/Self-Agency**	**Logical/Human Agency**	**Emotional/Human Agency**
I do not accept that offer.	That offer seems extremely unfair to me.	My client would not accept that offer.	That offer seems extremely unfair to my client.
I am not willing to pay that high of a price for so little in return.	Ouch, That offer would really hurt my wallet.	My client is not willing to pay that high of a price for so little in return.	Ouch, that offer would really hurt my client's wallet.
The phone is not worth that much to me.	This is going to make things really difficult for me.	The phone is not worth that much to my client.	This is going to make things really difficult for my client.
I would prefer a lower offer.	Wow. Can't you make me a more reasonable offer than that?	My client would prefer a lower offer.	Wow. Can't you make my client a more reasonable offer than that?
This offer does not make financial sense for me.	I simply cannot afford such a poor deal. It would leave me broke.	This offer does not make financial sense for my client.	My client simply cannot afford such a poor deal. It would leave them broke.
I have made some concessions, but I need you to make a better offer.	I have given in so much already. Aren't you willing to work with me a little on this?	I have made some concessions, but my client needs you to make a better offer.	I have given in so much already. Aren't you willing to work with my client a little on this?
Your offer needs to be lower if you want me to buy a phone.	I feel like you are trying to take advantage of me.	Your offer needs to be lower if you want my client to buy a phone.	I feel like you are trying to take advantage of my client.

Appendix D

HSIRB Approval Letter

Appendix D

HSIRB Approval Letter

WESTERN MICHIGAN UNIVERSITY

Human Subjects Institutional Review Board

Date: January 16, 2015

To: Chad Edwards, Principal Investigator
Brett Stoll, Student Investigator for thesis

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number 15-01-16

This letter will serve as confirmation that your research project titled "Human-Robot Negotiations" has been **approved** under the **exempt** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may **only** be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., ***you must request a post approval change to enroll subjects beyond the number stated in your application under "Number of subjects you want to complete the study***)." Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: January 15, 2016

1903 W. Michigan Ave., Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276
CAMPUS SITE: 251 W. Walwood Hall