THE ROLE OF THEORY OF MIND IN HUMAN-ROBOT INTERACTION

by

ISABELLA JARAMILLO

A thesis submitted in partial fulfillment of the requirements
For the Honors in the Major Program in Psychology
In the College of Sciences
and in the Burnett Honors College
at the University of Central Florida
Orlando, Florida

Summer term 2015

Thesis chair: Dr. Daniel S. McConnell

Abstract

Theory of Mind (ToM) has repeatedly been defined as the ability to understand that others believe their own things based on their own subjective interpretations and experiences, and that their thoughts are determined independently from your own. In this study, we wanted to see if individual differences in ToM are capable of causing different perceptions of an individual's interactions with human like robotics and highlight whether or not individual differences in ToM account for different levels of how individuals experience what is called the "Uncanny Valley phenomenon" and to see whether or not having a fully developed theory of mind is essential to the perception of the interaction. This was assessed by inquiring whether or not individuals with Autism Spectrum Disorder (ASD) perceive robotics and artificially intelligent technology in the same ways that typically developed individuals do; we focused on the growing use of social robotics in ASD therapies. Studies have indicated that differences of ToM exist between individuals with ASD and those who are typically developed. Comparably, we were also curious to see if differences in empathy levels also accounted for differences in ToM and thus a difference in the perceptions of human like robotics. A robotic image rating survey was administered to a group of University of central Florida students, as well as 2 surveys – the Autism Spectrum Quotient (ASQ) and the Basic Empathy Scale (BES), which helped optimize a measurement for theory of mind. Although the results of this study did not support the claim that individuals with ASD do not experience the uncanny valley differently than typically developed individuals, there were significant enough results to conclude that different levels of empathy may account for individual differences in the uncanny valley. People with low empathy seemed to have experienced less of an uncanny valley feeling, while people with higher recorded empathy showed to experience more of an uncanny valley sensitivity.

Dedication

To anyone who works with or works to help an autistic child in their lives on a daily basis, you are true superheroes.

And to my family:

To my sister, who never lets me forget that I am capable of accomplishing anything I set my mind to

To my brother, who has always looked out for me and always will

To my mother, who has always been my number one supporter and who has shown me the importance of going through life with a positive mind set

And to my father, without whom I would not know the meaning of perseverance and the value of a well-rounded education. I am the woman I have become because of you.

I love you.

Acknowledgements

I would like to express my gratitude to everyone who helped make this thesis possible. First and foremost, I would like to thank my thesis chair, Dr. Daniel S. McConnell. You have been an incredible professor and mentor over the last two years. Your guidance, time, and patience throughout my thesis is something I will never be able to thank you enough for, for without them, none of this would have been possible. Thank you for seeing the promise in me as a student and thank you for encouraging me to do this. To my committee members Dr. Matthew Chin and Dr. Jamie Schwartz, thank you both immensely for agreeing to be essential parts of my team and this project. I admire the both of you for your intellect and accomplishments, and am grateful for anything you contributed throughout this process. To Kelly Astro, Denise Crisafi, and anyone who is responsible for keeping the Honors in the Major program alive and well, I want to thank you for this amazing opportunity that so many students, including myself, are given because of you. To the University of Central Florida, for being a place where there are no limits to the goals I can set and the things I can accomplish. And to my parents, I would not be here receiving such a remarkable education if it weren't for you. To my brother and sister and the rest of my family, to all of my amazing friends, teachers, and to any employee or faculty member I have either had pleasure of knowing or working with over the last four years, thank you for everything that you have contributed to my life and to this moment. Thank you, from the bottom of my heart.

TABLE OF CONTENTS

LIST OF FIGURES

Section 1: Robot Human interaction and the Uncanny Valley

There has been a growing intrigue in studies relating to the interaction between robots/Artificially intelligent technologies over recent years. Our society has been progressing in a way that promotes the use of computerized and animatronic object to assist us in a numerous amount of daily tasks. It is not uncommon in this day and age to encounter a multitude of artificially intelligent (AI) objects, many of which have been accepted and even incorporated into daily life. Many modern devices that are automated or contain some form of AI have shown a great deal of positive response from society, mainly due to the convenience and ease that these objects continuously provide. A few examples of basic everyday encounters with AI include things such as the voice activation in cell phones (i.e., the popular use of SIRI on iPhones and Apple products) with text messaging assistance and automated responses, social media sites that use algorithms to learn your interests and preferences in order to display advertisements that will strike your appeal, to the multitude of technologies being incorporated into our schools, businesses, and schools, or even video games and voice activated game consoles. Aside from basic examples, there also exist entities that display a more complex form of AI, such as Cleverbot - a web application that uses an intelligent algorithm that allows it to have naturalistic conversations with humans by learning from the interaction and responses. As mentioned before, the ease that many of these technology have added to our daily lives has made it easy for people to accept them. According to Billings, Schaefer, Chen & Hancock (2012), one of the main components of a successful interactions between humans and robots come from the level of trust humans have in a particular technology to collaborate; trust that these systems will behave as they are expected to. In a time where robots are continuing to make a transition from being viewed as tools, to now being seen as a part of the team, trust will remain to be an essential factor in that relationship. The growth of this artificially intelligent world we are adapting will not be able to successfully thrive without trust in the systems we are choosing to accept.

It is however important, as convenient as they are, to remember that robots and technologies dot always work 100% of the time. It's hard to trust something that is expected to glitch or shut down for an unknown reason. On top of a looming threat of betrayal, there are also a few other negative aspects that have become apparent when dealing with AI and robotic technology. In the field of human-robot interaction, there exists a phenomenon termed the uncanny valley. While most of the population seems to have accepted many of the modern robotics and AI that has been incorporated into our daily lives, the continuous advancement of social robotics is something that should be addressed in terms of this phenomenon. In 1970, Masahiro Mori, a robotics professor at the Tokyo Institute of Robotics, wrote a paper that expressed his ideas in what he believed people's reactions would be to robots that appeared to be almost human. Mori's theory suggests that, as technology and robotics start to become more human-like, there comes a point where these artificial characters stop being so easily accepted and start becoming more unsettling to people; this area of uncertainty and sudden rejection is what Mori defined as the uncanny valley. Mori described the relationship of a human's affinity towards AI as a function of human likeness. As human likeness in a robot increases, so does a person's affinity towards the object. However, once the AI object reaches a specific point of human likeness, a sudden decrease in acceptance occurs between this almost human like entity, and the image of an actual healthy person. The sudden "dip" in the relationship that is illustrated by Figure 1 is a visual representation of the uncanny valley (Mori, 2012). Mori further explored this idea by explaining how people might feel when encountering a person with a very life-like prosthetic/robotic arm. His theory continued to explain that if the motion and appearance of a single prosthetic is capable of triggering negative reactions in people, then “a whole robot would magnify the creepiness” (Mori, 2012, p. 33-35).

McDorman (2006) further explored the uncanny valley phenomenon. In this study, 56 participants proceeded to rate 13 different robots and one human in scales from unfamiliar to familiar, not eerie to extremely eerie, and mechanical to human like. The results of this study suggested that human like-ness was not the only influence in the perceptions of the different constructs being measured in this experiment – rather the mechanical nature of the robots can also elicit negative reactions. Further, this study showed stronger negative responses to videos of moving robots compared to studies that typically used still images. The results from a more recent study (Jari Katsyri, Forger, Makarainen & Takala, 2015) suggested that while not all human-likeness manipulations may cause feelings of rejection and uncertainty, manipulated stimuli where the combination of robot and human-like features are inconsistent result in stronger uncanny responses (i.e. human eyes on an artificial face).

The uncanny valley has proved to be present in the realm of human-robot interaction, but it is also important to remember that while some people may experience feelings of unease and rejection, others maybe feel it to a lesser extent, or not at all, however, there is currently no research on identifying individual differences in terms of the uncanny valley.

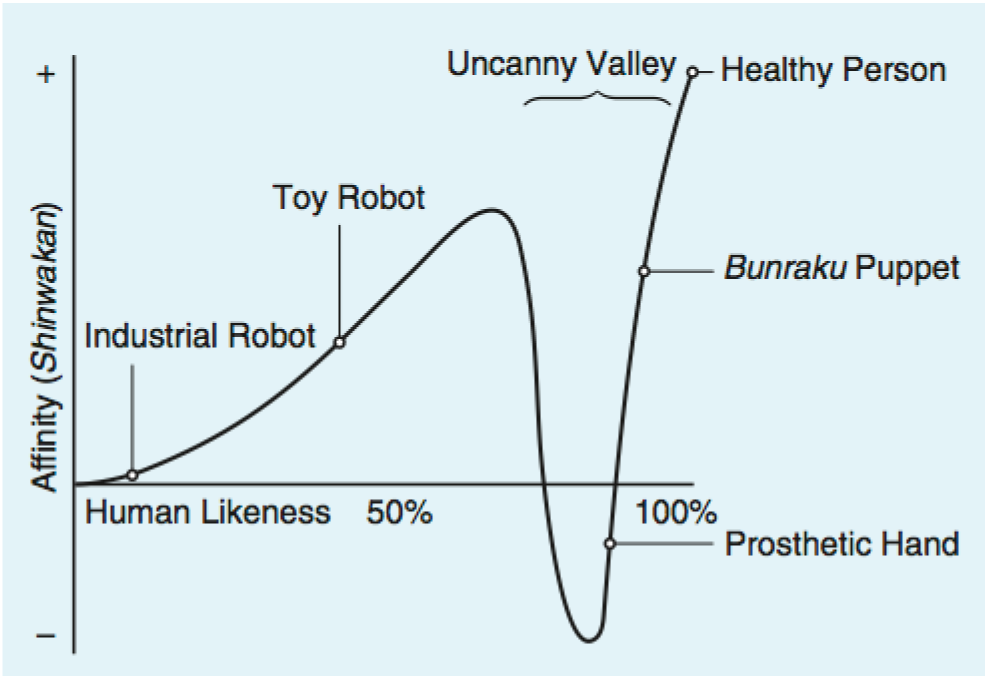


Figure 1 the Uncanny Valley (Mori 2012)

Section II: Theory of Mind and the Uncanny Valley

Theory of Mind (ToM) has been explained as the ability to represent, conceptualize and reason about different mental states (Schlinger 2009). A child typically develops both physically and mentally, it is natural for them to begin to understand how the world around them works. Alterations in ToM in a child become apparent around the age of four, when they begin to understand themselves and own mental thought processes, as well as gain an understanding that others are capable of going through their own independent mental processes. Furthermore, ToM has been suspected to underlie the causes of human behaviors, as well as the activity that is seen in one's conscious and subconscious (Schlinger 2009). Baron-Cohen, Leslie and Firth (1985) further continue to define ToM as "knowing that other people know, want, feel or believe things" (Baron-Cohen, et al., 1985, p. 38). It is further explained as the ability to attribute mental states to oneself and others and in turn be able to infer other people's belief's in regards to a specific situation, which would then allow an individual to predict what others will do or how they will react (Baron-Cohen, et al., 1985). Humans have been linked to a natural tendency of behaving toward others as if they possess intentions, or mind. Dennett (1987) outlines a theory that brings to light the idea that we may never really know whether other people minds and mental processes; weather this is true or not, we behave in a way in which we assume that they do. This inclination has been referred to as the intentional stance (Dennett 1987). Humans are naturally expected to attribute mind to other humans around them; but do these same principals apply when a person comes into contact with something that appears human but in fact isn't?

The previous section talked about the uncanny valley phenomenon, and how a feeling of uncertainty and unease is triggered in humans when they come into contact with human/life like robotics. Triggered by either human likeness, or a combination of both human and non-human characteristics, there can exist several different reasons as to why this feeling is actually manifested. It is possible that the uncanny valley feeling may be underlined by a feeling of mistrust in the opposing entity. As mentioned earlier by Billings et al. (2009) trust in the system is essential for a successful relationship between a human and a robot. The misleading human like appearance and mannerisms on something that is not entirely human can cause a feeling of distrust. The lack of consistency throughout the entity could be what causes the feeling of uncanny. Another related reason could be the fact that when a robot approaches some level of realism, it may lead to the tendency in others to try to attribute mind to the robot – that is, human like appearances are associated with the ability to think (as outlined by Dennett's (1987) intentional stance theory).

Section III: Autism Spectrum Disorder (ASD) and Differences in Theory of Mind

Autism spectrum disorder (ASD) is a term used to describe a series of complex disorders that occur during the earlier stages of brain development. Classified as a neurodevelopmental condition, the curious nature of this disorder has made it a widespread topic of interest and research in recent years. Although it is linked to neurodevelopmental deficiencies as wells as resilient underlying genetic factors, knowledge on the exact etiology of ASD is still not entirely known (Myers & Johnson, 2007).

Autism is marked by social and cognitive impairments that are often accompanied by abnormal displays of physical behavior or certain learning disabilities (American Psychiatric Association (APA) 2013). The signs of autism can begin to be seen as early as six months of age and are likely to become more apparent from ages two to three. When a child/individual is diagnosed with autism, several factors are taken into consideration. physicians and psychologists experts focus on the level of difficulty in social interactions, how an individual communicates- whether it's verbal or nonverbal, and whether or not the individual demonstrates any harmful, repetitive or compulsively unconventional behavior (Myers & Johnson, 2007). Because autism is not limited to just one specific classification, there exists an autism spectrum, which helps categorize each specific disorder, which has shown to be specific/unique to each individual. The Diagnostic Statistical Manual of Mental Disorders, Fifth Edition (DSM-V), is the guideline that is currently used by health professionals in the diagnosis of ASD. The DSM-V outlines series of general categories, such as communication and social interactions, and arrives at a diagnosis by specifically pinpointing certain abnormalities within each category that would indicate if an individual falls somewhere on the autism spectrum.

There have been studies that highlight apparent individual differences in theory of mind between individuals with ASD and typically developed individuals. ASD is often mainly distinguished by its significant deficits in communication and social interactions; this includes deficits in social exchanges, nonverbal communication and the ability to cultivate the necessary skills needed to develop, maintain and understand relationships (APA 2013). Baron-Cohen, Leslie and Firth (1985) proposed an idea that explained these pervasive social impairments - the notion that children with autism lack the development of second-order representations, and therefore lack a crucial aspect of social skills described by the term 'theory of mind' The ability to assume behavioral outcomes in another individual could be considered an essential skill that promotes appropriate social interaction and his idea inherently links the lack of a ToM to the impairment seen in a child with autism's social skills. In addition, if lacking a ToM is characterized as an underlying factor of an individual with autism's underdeveloped social skills, then it can also be associated with an individual with autism's inability to engage in appropriate social exchanges and can be assumed that individuals with autism are also incapable of showing appropriate levels of empathy.

In their experiment, Baron-Cohen et al. (1985) examined 20 diagnosed children with autism as per the established criteria, 14 children with Down syndrome, and 27 typically developing preschool children, all of similar mental ages. The experiment involved a method now known as the Sally-Ann test. The experimenter tells the child a story, in which a doll named Sally places a marble in her basket, and then briefly leaves the room. In the time of the first doll's absence, a second doll, named Ann, takes the marble out of the basket and places it into her box. When Sally returns, the experimenter asks the child the question "where will the first doll look for the marble?" If the child answers with the marble's new location, they fail to account for the first dolls belief of the marble's location. The results of this experiment show that, unlike the groups of children with Down Syndrome and typically developing preschool children, who pointed to where the marble was in the first place (which was the response that the researchers were looking for), the group of children with diagnosed autism consistently pointed to where the marble was relocated. Failure of the children with autism to provide the desired answer was neither attributed to random pointing or memory impairments. Rather, it was concluded that the children with autism merely did not accept the difference between their own and the doll's knowledge. These results suggested that the group of children with autism failed to exhibit ToM.

Up until the late 1990's, studies that focused on testing ToM all concurrently hit a limit where the only participants being tested were those of a mental age of about 6. Since these tests only seemed to be administered to children or those with the mental age of children, it would have been wrong to assume that all individuals with autism lacked a ToM ability. It made sense to administer such tests to individuals with the developmental age of an adult, who still exhibited some form of ASD, to test if the results of the studies conducted with children applied to anyone with ASD. Thus, a comparable study attempted to test for ToM on higher functioning adults diagnosed with autism (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). In this study, three groups of adults were tested; 16 participants with ASD, 50 typically functioning adults, and 10 adults with Tourette's syndrome. All groups were exposed to the same "Eye Task" test. This particular test consisted of 25 different pictures of faces. Each picture only showed a portion of each face, from midway along the nose to a little bit above the eye brows (just enough to see the facial expression in the persons eyes). Each participant was shown each picture for approximately 3 seconds and were asked to pick, from two words that were given, which one best described the expression on the face. For example, a picture would be shown and the participant would be asked to pick between the words “concerned” or “unconcerned” to describe the face. Two controls of gender recognition and basic emotion recognition were also administered to all of the participants to make sure that the deficits of the eye task were not due to alternative factors. The results of this study indicated that adults with autism also demonstrated impaired ToM.

Section IV: Social Robotics and ASD Therapies

If an autistic child’s inability to develop a theory of mind is said to cause such severe impairment in social skills, then how does this reflect an autistic individual’s ability to interact with modern technology? It is not uncommon in this day and age to encounter a multitude of artificially intelligent objects, much of which have been accepted and even incorporated into daily life. Many modern devices that are automated or contain some form of artificial intelligence (AI) have shown a great deal of positive response from society, mainly due to the convenience and ease that these objects continuously provide. A few examples of basic everyday encounters with AI include things such as the voice activation in cell phones (i.e. the popular use of SIRI on iPhones and Apple products) with automated responses, social media sites that use algorithms to learn your interests and preferences in order display advertisements that will strike your appeal, or even video games and voice activated game consoles. Aside from basic examples, there also exist entities that display a more complex form of AI, such as Cleverbot- a web application that uses an intelligent algorithm that allows it to have naturalistic conversations with humans by learning from the interaction and responses.

It is important to inquire whether or not an autistic individual is able to appropriately interact with AI because, in recent years, a major developing application for social robotics has been incorporating it into certain therapies that gear towards assisting children with ASD (Cabibihan, Javed, Ang Jr., & Aljunied 2013). Through online interactive games and similarly engaging activities, these socially interactive robots are used to relate with these children, by creating enticing and relatable situations, making a child feel compelled to interact. The main goal is to target the technology behind robotics towards training these children in certain skills that provoke desired behaviors through a system that displays encouragement as well as positive feedback when a task is completed without fail (Cabibihan et al. 2013). One of the most promising applications of this technology is using these social robotics as an alternative way for a child to practice “context-appropriate” social skills; because an autistic child’s social skills are underdeveloped, these robots serve as a chance for a child with ASD to exercise appropriate behaviors through indirect experiences (Cabibihan et al. 2013). Additional advantages to incorporating social robotics in a clinical setting include the fact that robots can be programmed to adapt and allow individualized treatment specific to each child. Furthermore, there appears to be an inherent appeal to robotics among the individuals who lie within the autism spectrum (Diehl, Schmitt, Villano & Crowell, 2012). Welch, Lahiri, Warren, and Sarkar (2010) showed that virtual robots were judged to be effective as experimental social stimuli for children with ASD and concluded that this type of research may be able to develop a better understanding of the susceptibilities of social communication in children with ASD. Robots can be programmed in a way that allows their interactions to generate a response from a child that promotes positive prosocial behavior. This may even be useful for promoting enjoyment in social activities for children with ASD (Diehl, Schmitt, Villano & Crowell, 2012). Furthermore, research suggests that robots can potentially be used as a diagnostic tool. A robot could be designed in a way that would provoke a set of social responses in reaction to a set of social pressures. The diagnosis would be attained by measuring the presence, absence, or quality of the child's overall response. Scassellati (2007) also proposed the idea of using social robots as a diagnostic tool. With a wide range of anthropomorphic characteristics, these robots are expected to generate positive responses from children in terms of motivation and engagement (Scassellati, 2007). Although there are extensive studies that explore the effects of certain robotic interaction levels on typically developed adults, "very little is known about how individuals with autism respond to these design dimensions" (Scassellati, 2007, p. 3).

With so much potential for the incorporation of robotics and AI in ASD therapy it is, again, important to inquire about the interactions between these children and these robotic technologies. Do these interactions entail the same type of social psychological mechanisms as person-to-person interaction? Do the children engage in social perception and cognition toward social robots? And lastly, if there is such a thing as differences in ToM or lack thereof, what role does it play in human-robot interaction and account for differences in the uncanny valley experience?

Section V: Hypothesis

Are individual differences in theory of mind capable of causing individual differences in the experience of the uncanny valley? Although there is little to no research verifying that individual differences in the uncanny valley exist, perhaps differing states of ToM is an underlying aspect that should be considered.

It is apparent that studies have suggested differences in ToM between individuals with ASD and typically developed individuals. So, what occurs when an autistic individual, said to lack a properly developed theory of mind, interacts with the same entities that give typically developed humans that uncanny feeling? If a fully developed theory of mind is a crucial part of what allows us to make mental attributions of these robots, and if the Uncanny Valley is a result of attributing mind to these realistic entities, and if the "underdeveloped theory of mind" idea is a valid explanation for the setbacks seen in autism, then it can be predicted that those with ASD will not experience the Uncanny Valley phenomenon. Similar results are expected if we measure a similar construct such as empathy. If empathy proves to be a valid component of the emotion that is experienced through the uncanny valley, then we can also predict that those with low empathy will also be unable to experience the uncanny valley phenomenon, or at least react to a lesser degree.

As a preliminary investigation of this hypothesis, we will examine healthy adults' reactions to potentially uncanny stimuli. While we will not study those with ASD, we will measure the participants with two scales. The Autism Spectrum Quotient (ASQ) will be used to measure autistic-like traits in the group of participants, and the Basic empathy scale (BES) will be used to separately measure the levels of empathy in the same group of participants. If differences in ToM appear in those with ASD and those without, then it is also possible that another related concept, such as empathy, may also account for individual differences in ToM and there for individual differences in the uncanny valley.

We hypothesize that those who score high on the ASQ will experience the uncanny valley to a lesser degree than those who score low on the ASQ. Likewise, those who score high on the BES will be expected to reflect a higher sensitivity to the uncanny valley than those who exhibit lower empathy scores.

Method

Participants

Participants (N = 217) were recruited how, need to describe from a large southeastern university in the US using an online recruitment system managed by the university. Participants were undergraduate students enrolled in psychology courses and received extra or partial course credit in exchange for participation. All participants provided consent in which they agreed to all of the terms and conditions of the study via a form that was electronically administered to each participant prior to beginning the actual survey. All protocols were approved by the university Institutional Review Board (IRB).

Materials

Participants viewed and rated 40 images of robots, ranging from friendly non-humanlike robots to various humanlike or unfriendly robots (See Appendix A). These images were randomly selected from an internet search that aimed to collect a wide range of robotic images and representations. Each image was rated according to these 5 questions: 1) I would feel comfortable being near this robot, 2) I would feel comfortable interacting with this robot, 3) I would dread it if I knew I had to meet this robot, 4) I feel positive feelings about this robot, and 5) I think it might be fun to interact with this robot. For each individual image, these five questions were displayed by a matrix which allowed each of them to be ranked on a 5-point likert scale; the five ranks range from agree, strongly agree, neutral, disagree and strongly disagree (see Appendix B). Each question was capable of earning a total of 5 points – because of the negative nature of the third question in comparison to the others, each third question for each image was reverse coded in order to be scored in the same manner as the other four questions. Each question having a total worth for 5 points made for a possible sum of 25 points for the total ranking of each image.

Participants completed the 20-item Basic Empathy Scale (BES) (Jolliffe & Farrington, 2006) (see Appendix D). This test served as an operational definition of theory of mind, as it measured the participants' abilities to both experience and comprehend the emotional states of others. Participants also completed the 50-item Autism-Spectrum Quotient (ASQ) (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) (see Appendix C). While not diagnostic of ASD, it has been shown to detect autism-like traits in the normal population. All materials were presented using an online survey management platform managed by the university.

Procedure

Participants who volunteered for the experiment using the recruitment system were provided a link to the survey. After providing consent, the participants viewed each of the 40 images, one per page, in a random order. Each image was accompanied by the five rating questions. Subsequent to the image ratings, participants completed the BES and ASQ, in that order, and were thanked for their participation.

Results

Of the 217 participants who attempted the survey, several were removed. One participant failed to complete the ASQ and also skipped two images without rating them. Six participants completed the entire survey in under 10 minutes. In comparison to the average completion time of approximately 18 minutes, we judged that completion times under 10 minutes reflected participants who hurried through the survey and were unlikely to have given honest and thoughtful responses. Data from the remaining 210 participants were thus included in the subsequent analyses. The BES and ASQ scales were scored according to the original authors' instructions. The data that was collected for the ASQ displayed a near normal distribution across participants (M=18.79, SD=4.97) with a range of 28.00. The data from the BES also resulted in a near normal distribution (M=75.91, SD=8.94) with a range of 47.00. Based on the data collected from the BES and the ASQ, the participants were divided into groups based on their scores. There was a significant negative correlation between the BES and ASQ scales, $r = -.18$, $p = .008$. The scores for both questionnaires were ranked, and the top 1/4th and the bottom 1/4th percentiles for both surveys were then determined. The lower quarter of the BES cut off at a score of 69 and the upper quarter cut off at a score of 81. Likewise, the lower quarter of the ASQ cut off at a score of 15 and the upper quarter cut off at a score of 22. This led to the creation of four different groups, low ($n = 54$) and high ($n = 58$) empathy, and low ($n = 56$) and high ($n = 60$) ASQ score. Participants whose responses fell between the low and high Cutoff scores for either group were excluded from the analyses.

The sums of the ratings of the robot image were determined and the averages for each photo were calculated from the sums. The average rating for each image was compared to the scores that each participant received on both the ASQ and the BES. When comparing image ratings in terms of the low and high ASQ groups, there were no significant differences. The graph shown in Figure 2 shows a linear distribution for the means in both groups. The linear fit regression lines for both groups are almost equal in slope.

However, when the mean image ratings of the two BES groups were compared, a significant difference was apparent. Those whose scores reflected high empathy found some of the images significantly more unpleasant than those with lower empathy scores. The regression line for the group with higher empathy shows a steeper slope than that of the low empathy group. This relationship can be seen in the graph in Figure 3. In addition, two-tailed independent group t-tests were performed between several of the images between the low and high BES groups. Using a corrected alpha ($p < .01$) due to multiple comparisons, significant differences were found for images 6, 22, 47, and 52.

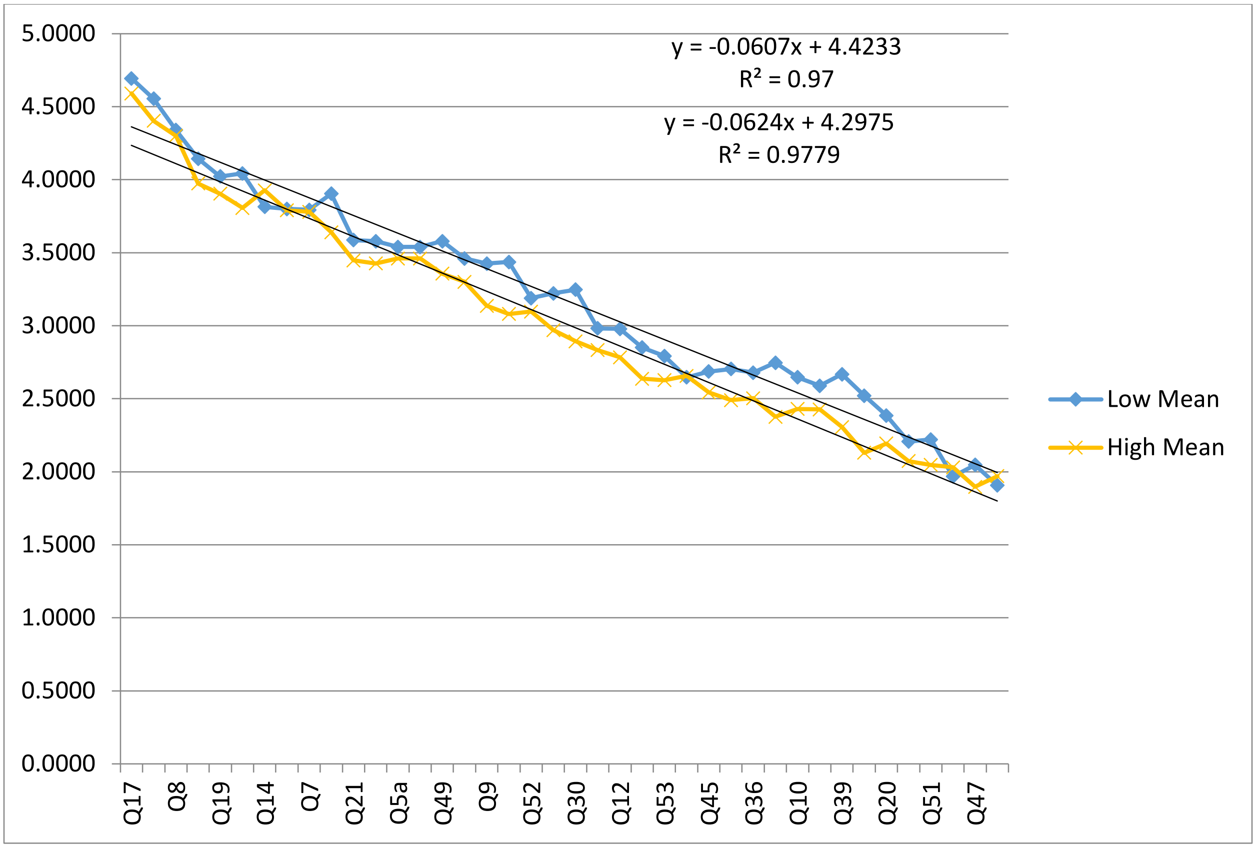


Figure 2 Mean image ratings for each image, sorted from highest to lowest mean overall rating. Separate lines indicate the lowest and highest scores on the ASQ scale.

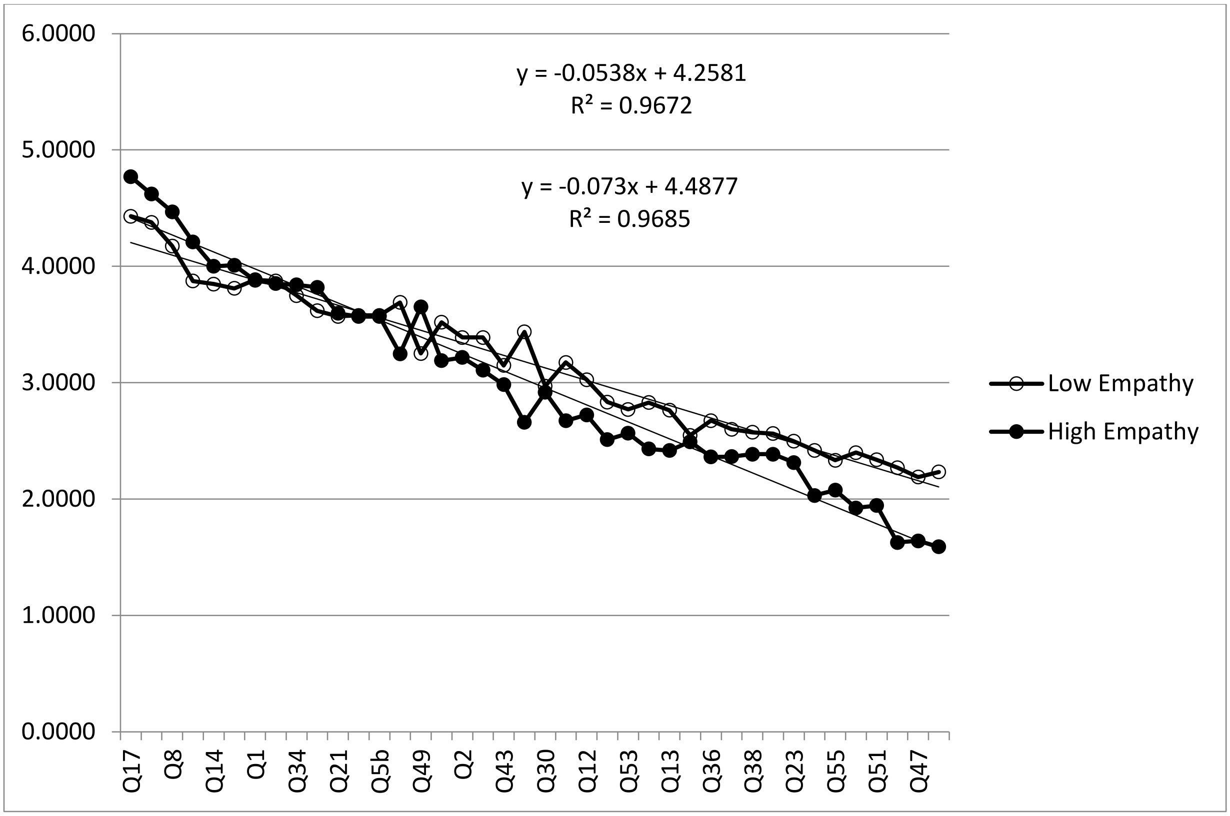


Figure 3 Mean image ratings for each image, sorted from highest to lowest mean overall rating, with separate lines for low and high BES groups.

Discussion

The results found in this study were not enough to support the claim that autistic individuals do not experience the uncanny valley, though we have some evidence that theory of mind may be linked to the uncanny valley.

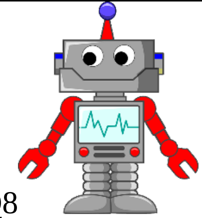
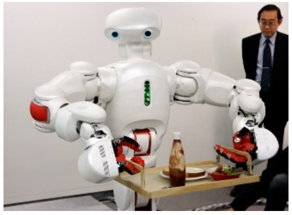
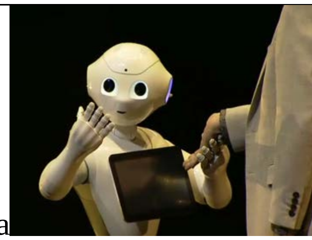
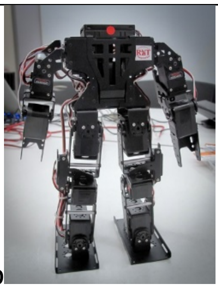
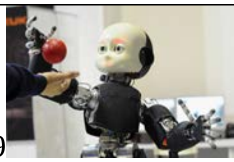
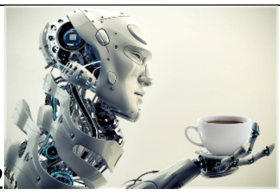
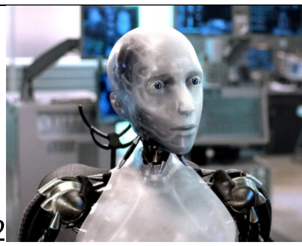
We used the ASQ as an artificial way to create an operational "autistic" population for our study, and failed to find differences in negative reactions to robots (or robotic images in the case of this experiment) as a function of those who scored high and low on this scale. While the ASQ may detect certain autism-like traits, it may not be an accurate test of ToM or ASD, and thus failed to detect characteristics relevant to the social perception of robots. Alternatively, it may be the case that autistic characteristics are not as directly linked to ToM as

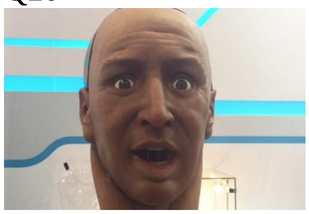
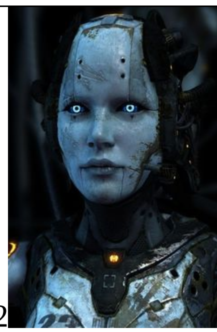
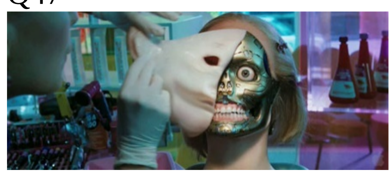
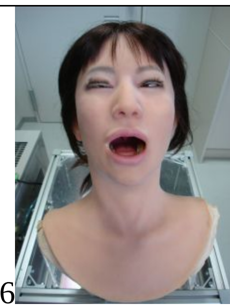
Baron-Cohen, et al. (1985) originally predicted. This is, of course, assuming that the uncanny valley is a product of a properly developed ToM. These results may also be explained by the fact that we did not directly test a group of ASD diagnosed individuals; it is possible that different results may become apparent if this hypothesis was tested on actual autistic individuals. This type of work is necessary to clarify the interpretations of our negative findings regarding the ASQ.

We did, however, find significant results when comparing the individual image ratings to the BES. There was a significant difference in how some of these images were perceived between the groups of participants who had high empathy scores, and those who had low empathy scores. Those with higher recorded empathy perceived some of the more unpleasant looking images more negatively than those with low empathy scores. It's understandable that the uncanny valley responds positively when compared to levels of empathy. It is possible that the uncomfortable feeling one goes through when experiencing the uncanny valley is linked to emotional responses- certain images will either make you feel positive emotions or negative emotions towards them. To the extent that empathy may be a component of theory of mind, we thus have tentative support for the idea that theory of mind is part of the uncanny response.

As previously mentioned, social robotics are being used in therapies that can help enhance an autistic child's social skills and serve as a way to practice using appropriate social interaction tactics. Research on social perception of robots has shown that people can respond appropriately to different social cues exhibited by robots (Fiore, Wiltshire, Lobato, Jentsch, & Axelrod, 2013). However, these studies have not taken into account individual differences in terms of empathy or theory of mind, and our findings suggest that continued work in social robots must incorporate these differences, and seek to clarify other factors that may influence social perceptions. Typically developed individuals are usually aware of their interactions with human-like robotics and are aware of the social cues that trigger their perceptions of said interactions. The significant scores shown through comparison to the Basic Empathy Scale highlight a link between empathy, and the perceptions of social cues from a robot. It may be likely that when studies are done on individuals and their perceptions of robots, the levels of empathy in each participant might also need to be taken into consideration.

APPENDIX A: Images shown on the Individual rating scales, organized by the averages of the scores they were given, from highest score (more positively perceived) to lowest score (more negatively perceived)

Q17 	Q37 	Q8 	Q3 
Q14 	Q19 	Q6 	Q7 
Q34 	Q18 	Q21 	Q5a 
Q5b 	Q32 	Q 49 	Q9 
Q2 	Q15 	Q43 	Q52 
Q30 	Q11 	Q12 	Q16 
Q53 	Q10 	Q13 	Q41 

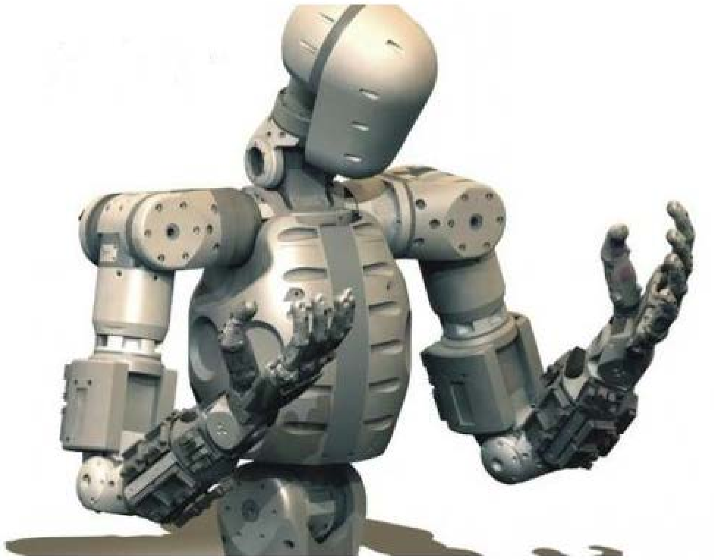
Q36 	Q45 	Q38 	Q39 
Q23 	Q20 	Q55 	Q4 
Q51 	Q22 	Q47 	Q6 

Images used by the researchers in the image rating survey

APPENDIX B: sample of the scoring system for each individual image

Page Break

Q15



	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I would feel comfortable being near this robot.	○	○	○	○	○
I would feel comfortable interacting with this robot	○	○	○	○	○
I would dread it if I knew I had to meet this robot	○	○	○	○	○
I feel positive feelings about this robot	○	○	○	○	○
I think it might be fun to interact with this robot	○	○	○	○	○

Page Break

Q16

Store

APPENDIX C: The Autism Spectrum Quotient (ASQ)

The Autistic-Spectrum Quotient

Note: Answer choices for all items is:

definitely agree	slightly agree	slightly disagree	definitely disagree

1. I prefer to do things with others rather than on my own.
2. I prefer to do things the same way over and over again.
3. If I try to imagine something, I find it very easy to create a picture in my mind.
4. I frequently get so strongly absorbed in one thing that I lose sight of other things.
5. I often notice small sounds when others do not.
6. I usually notice car number plates or similar strings of information.
7. Other people frequently tell me that what I've said is impolite, even though I think it is polite.
8. When I'm reading a story, I can easily imagine what the characters might look like.
9. I am fascinated by dates.
10. In a social group, I can easily keep track of several different people's conversations.
11. I find social situations easy.
12. I tend to notice details that others do not.
13. I would rather go to a library than a party.
14. I find making up stories easy.
15. I find myself drawn more strongly to people than to things.
16. I tend to have very strong interests, which I get upset about if I can't pursue.
17. I enjoy social chit-chat.
18. When I talk, it isn't always easy for others to get a word in edgeways.
19. I am fascinated by numbers.
20. When I'm reading a story, I find it difficult to work out the characters' intentions.
21. I don't particularly enjoy reading fiction.
22. I find it hard to make new friends.
23. I notice patterns in things all the time.
24. I would rather go to the theatre than a museum.
25. It does not upset me if my daily routine is disturbed.

26. I frequently find that I don’t know how to keep
a conversation going.
27. I find it easy to “read between the lines” when
someone is talking to me.
28. I usually concentrate more on the whole picture,
rather than the small details.
29. I am not very good at remembering phone numbers.
30. I don’t usually notice small changes in a situation,
or a person’s appearance.
31. I know how to tell if someone listening to me
is getting bored.
32. I find it easy to do more than one thing at once.
33. When I talk on the phone, I’m not sure when it’s
my turn to speak.
34. I enjoy doing things spontaneously.
35. I am often the last to understand the point of a joke.
36. I find it easy to work out what someone is
thinking or feeling just by looking at their face.
37. If there is an interruption, I can switch back to
what I was doing very quickly.
38. I am good at social chit-chat.
39. People often tell me that I keep going on and
on about the same thing.
40. When I was young, I used to enjoy playing games
involving pretending with other children.
41. I like to collect information about categories
of things (e.g. types of car, types of bird,
types of train, types of plant, etc.).
42. I find it difficult to imagine what it would be
like to be someone else.
43. I like to plan any activities I participate in carefully.
44. I enjoy social occasions.
45. I find it difficult to work out people’s intentions.
46. New situations make me anxious.
47. I enjoy meeting new people.
48. I am a good diplomat.
49. I am not very good at remembering
people’s date of birth.
50. I find it very easy to play games with
children that involve pretending.

APPENDIX D: Basic Empathy Scale (BES)

BASIC EMPATHY SCALE

The following are characteristics that may or may not apply to you. Please circle one answer for each statement to indicate how much you agree or disagree. Please answer as honestly as you can.

1. My friend's emotions don't affect me much.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

2. After being with a friend who is sad about something, I usually feel sad.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

3. I can understand my friend's happiness when she/he does well at something.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

4. I get frightened when I watch characters in a good scary movie.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

5. I get caught up in other people's feelings easily.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

6. I find it hard to know when my friends are frightened.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

7. I don't become sad when I see other people crying.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

8. Other people's feelings don't bother me at all.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

9. When someone is feeling 'down' I can usually understand how they feel.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

10. I can usually figure out when my friends are scared.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

11. I often become sad when watching sad things on TV or in films.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

12. I can often understand how people are feeling even before they tell me.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

13. Seeing a person who has been angered has no effect on my feelings.

STRONGLY DISAGREE	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGREE**	**STRONGLY AGREE**

14. I can usually figure out when people are cheerful.

STRONGLY	**DISAGREE**	**NEITHER AGREE NOR**	**AGREE**	**STRONGLY**

DISAGRE E		**DISAGREE**		**AGREE**

15. I tend to feel scared when I am with friends who are afraid.

STRONGL Y DISAGRE E	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGRE E**	**STRONG LY AGREE**

16. I can usually realize quickly when a friend is angry.

STRONGL Y DISAGRE E	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGRE E**	**STRONG LY AGREE**

17. I often get swept up in my friend’s feelings.

STRONGL Y DISAGRE E	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGRE E**	**STRONG LY AGREE**

18. My friend’s unhappiness doesn’t make me feel anything.

STRONGL Y DISAGRE E	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGRE E**	**STRONG LY AGREE**

19. I am not usually aware of my friend’s feelings

STRONGL Y DISAGRE E	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGRE E**	**STRONG LY AGREE**

20. I have trouble figuring out when my friends are happy.

STRONGL Y DISAGRE E	**DISAGREE**	**NEITHER AGREE NOR DISAGREE**	**AGRE E**	**STRONG LY AGREE**

References

American Psychiatric Association (APA) (2013). *Diagnostic and Statistical Manual of Mental Disorders* (*5th ed*.) (DSM-V). Arlington, VA. American Psychiatric Publishing

Baron-Cohen, S., Jollidde, T., Mortimore, C., Robertson, M. (1997). Another Advanced Test of Theory of Mind: Evidence from Very High Functioning Adults with Autism or Asperger Syndrome, *Association for Child Psychology and Psychiatry, 38(7),* 813-822

Baron-Cohen, S., Leslie, A.M. & Firth, U. (1985). Does an autistic child have a "theory of mind"?. *Cognition, 21*, 37-46.

Baron-Cohen, s., Wheelwrite, S., Skinner, R., Martin, J. & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17

Billings D.R., Schaefer K.E., Chen, J.Y.C., Hancock P.A. (2012). Human-robot interaction: developing trust in robots, *HRI'12 proceedings of the seventh annual AMC/IEEE international conference on Human-Robot Interaction.* 109-110. doi: 10.1145/2157689.2157709

Cabibihan, J.J., Javed, H., Ang Jr, M. & Aljunied, S.M. (2013). Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *International Journal of Social Robotics, 5*(4), 593-618. doi 10.1007/s12369-013-0202-2

Dennett D.C. (1989). The intentional stance. Cambridge, Massachusetts. The MIT Press.

Diehl, J.J., Schmitt, L.M., Villano, M. & Crowell, C.R. (2012). The clinical use of robots for individuals with Autism Spectrum Disorders:A critical review. *Research in Autism Spectrum Disorders, 6*, 249-262. doi:10.1016/j.rasd.2011.05.006

Fiore, S.M., Wiltshire, T.J., Lobato E.J.C., Jentsch, F.G., Huang, W.H., Axelrod, B. (2013). Toward understanding social cues and signals in human-robot interaction: effects of robot Gaze and proxemics behavior, *Frontiers in Psychology, 4:859,* 1-15. doi:10.3389/fpsyg.2013.00859

Jolliffe, D., Farrington, D.P. (2006). Development and Validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*, 589-611

Katsyri, J., Forger, K., Makarainen, M. & Takala, T. (2015). A review of evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness, *Frontiers in Psychology*, 6(390), doi: 10.3389/fpsyg.2015.0039

Masahiro, M. (2012). The Uncanny Valley (K.F MacDorman., N. Kageki, Trans). *IEEE Spectrum*, [*"The Uncanny Valley," Energy, 7(4), pp. 33–35, 1970 (in Japanese)] (origioal work published in 1970)*

MacDorman, K. F. (2006) Subjective ratings of robots video clips for human likeness, familiarity, and eeriness: Am exploration of the uncanny valley. *ICCS/cogSci-2006 long symposium: Toward Social Mechanisms of android science,* 26-29

Myers, M.S., Johnson C.P, Med. & the Council on Children with Disabilities. (2007) Management of Children with Autism Spectrum Disorders. *American Academy of Pediatrics*, *120*(5), 1162-1182. doi:10.1542/peds.2007-2362

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

Schlinger Jr H.D. (2009). Theory of Mind: An Overview and Behavioral perspective, *The Psychological Record, 59,* 435-448

Welch, K.C., Lahiri, U., Warren, Z., Sarkar, N., (2010). An Approach to the Design of Socially Acceptable Robots for Children with Autism Spectrum Disorders. *International Journal of Social Robotics*, *2*, 391-403. doi: 10.1007/s12369-010-0063-x