Title
Game Theoretic Investigation of Decision-Making and Theory of Mind in Neurotypical Individuals with Differential Levels of Autistic Traits

Author
Craig, Alexis

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Game Theoretic Investigation of Decision-Making and Theory of Mind in Neurotypical Individuals with Differential Levels of Autistic Traits

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in Psychology - Cognitive Neuroscience

by

Alexis B. Craig

Dissertation Committee:
Professor Jeffrey L. Krichmar, Chair
Associate Professor Emily D. Grossman
Professor Wendy A. Goldberg

2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Thank you to my advisors, Jeffrey Krichmar and Emily Grossman, for your time, patience, and exemplary guidance over the course of my graduate education. Thank you to my committee member, Wendy Goldberg, for your input and expertise. Thank you to my family for your support and for giving me direction in life. Thank you to the members of the Cognitive Anteater Robotics Laboratory for your camaraderie and insight over the past five years.

CURRICULUM VITAE

Alexis Craig

EDUCATION

University of California, Irvine

Doctor of Philosophy, Psychology - Cognitive Neuroscience, 2016

Master of Science (MS), Cognitive Neuroscience, 2014

University of California, Los Angeles

Bachelor of Science, Cognitive Science, 2011
Minor in Human Complex Systems

PUBLICATIONS

Craig, A.B., Grossman, E.D., & Krichmar, J.L. (In preparation). Investigation of autistic traits through strategic decision-making in games with adaptive agents.

Craig, A.B., Phillips, M.E., Zaldivar, A., Bhattacharyya, R., & Krichmar, J.L. (2016). Investigation of biases and compensatory strategies using a probabilistic variant of the Wisconsin Card Sort Test. *Frontiers in Psychology*, *7*.

Asher, D.E., **Craig, A.B.,** Zaldivar, A., Brewer, A.A., & Krichmar, J.L. (2013). A dynamic, embodied paradigm to investigate the role of serotonin in cost and decision-making. *Frontiers In Integrative Neuroscience, 7*.

Craig, A.B., Asher, D.E., Oros, N., Brewer, A.A., & Krichmar, J.K. (2013). Social contracts and human-computer interaction with simulated adapting agents. *Adaptive Behavior, 21* (5), 371-387.

Cha, Y.H., Chakrapani, S., **Craig, A.,** & Baloh, R.W. (2012). Metabolic and functional connectivity changes in Mal de Debarquement syndrome. *PLoS ONE, 7* (11). doi:10.1371/journal.pone.0049560

Brewer, A.A., Asher, D.E., **Craig, A.B.,** Oros, N., & Krichmar, J.L. (2012). A Dynamic and Embodied Environment to Probe the Neural Correlates of Decision-Making and Social Signaling. Poster presented at: Society for Neuroscience: Dynamical Neuroscience XX: Collective Cognition - The Neurophysiology of Social Neuroscience (New Orleans, LA).

AWARDS AND DISTINCTIONS

Department of Social Sciences Associate Dean's Fellowship, 2013 – 2016

Brython Davis Fellowship, 2013

UCLA Latin Honors: Cum Laude, 2011

UCLA Dean's Honors List, 2007-2011

Tailhook LCDR Chris "Basher" Blaschum Memorial Scholarship, 2007

ABSTRACT OF THE DISSERTATION

Game Theoretic Investigation of Decision-Making and Theory of Mind in Neurotypical Individuals with Differential Levels of Autistic Traits

by

Alexis Craig

Doctor of Philosophy in Psychology – Cognitive Neuroscience

University of California, Irvine, 2016

Professor Jeffrey L. Krichmar, Chair

Theory of mind (ToM) is the cognitive ability to imagine the thoughts, beliefs, goals, and motivations of another person. This ability is utilized both consciously and subconsciously in the majority of social interaction in order to conduct rational, appropriate decision-making and form cooperative or competitive relationships to achieve one's goals. Automatic, or implicit, ToM is a trait that has been shown to be impaired in individuals who suffer from autism spectrum disorders (ASD), which is one of the factors leading to the social deficits experienced by these individuals. The objectives of this dissertation are to 1) develop a non-verbal social task in which implicit ToM is evoked; 2) develop a computer agent that is capable of acting as a social partner within the task to evoke ToM response; 3) use this task in behavioral research to collect data from neurotypical individuals who engage socially with the computer agent. The first chapter introduces the Stag Hunt, a game theoretic task with a payoff matrix biased toward cooperation in which two players must decide whether to hunt a low payoff hare individually or attempt to cooperate to catch a high payoff stag. For the purposes of conducting the task with subjects, an adaptive agent was developed that initially enters the task naïve and develops in real time, based on the actions of the other player and the outcomes of repeated trials, a strategy suited to the subject it is interacting with.

The second chapter discusses a behavioral study using a probabilistic variant of the Wisconsin Card Sorting Test in order to investigate cognitive biases inherent in decision-making in an uncertain environment. Understanding common coping mechanisms for uncertainty provides a baseline for comparing the decision-making strategies of atypical individuals. The third chapter expands upon the efforts of the first with the inclusion of variants of the preexisting adaptive agent in order to differentially evoke complex ToM responses correlated to the level of autistic traits in subjects drawn from the general population. This was accomplished through use of forward planning and simulated ToM. The fourth chapter discusses a pilot fMRI study to collect data on the correlation between levels of autistic traits and brain activation related to ToM and decision-making. Taken together, this body of work provides a foundation for utilizing non-verbal social tasks to evoke ToM response with non-human agents, a format that lends itself well to autism research. The overarching goal of this line of work is to aid in the identification of differences in neural processing of individuals affected with varying levels of ASD, both clinically and subclinically, to provide information that can be traced back to the locus of development in the brain leading to autistic expression.

INTRODUCTION

On a day-to-day basis, living things utilize cooperation and competition to reach a desired outcome. Because of this common dynamic, social behavior in cooperative and competitive situations has become a popular field of study. The use of games in social behavior experiments can give insight into the interactive dynamics between players, as well as their decision-making processes. Such games can highlight individual and group differences in a controlled and highly customizable environment. Game theory provides additional benefits, as it includes tools to predict behavior and decision-making by assuming players will attempt to achieve the most desirable outcome (Lee, 2008). Games are especially useful when considering the topic of social behavior from a human-computer interaction (HCI) standpoint. Because games provide a clearly defined state space and set of rules, they are amenable to providing a framework for humans to interact with computers as partners or opponents. The Prisoner's Dilemma, Ultimatum Game, Trust Game, Hawk/Dove, and Stag Hunt are among the most prominent games used to research social behavior in HCI.

Game theory and computer agents

In a study conducted by Kiesler, Sproull, and Waters (1996), the Prisoner's Dilemma was used to determine the differences in cooperation between humans and different types of computer opponents. In the Prisoner's Dilemma, two players must decide to either "rat out" their opponent or to keep quiet, a decision that affects each player's "sentencing," or personal cost. In these experiments, subjects played against three types of computer opponents: text-based, electronically generated speech-based, and electronically generated face- and speech-based. The text-based opponent interacted with the human player through text messages, while the speech-based opponents used computer-generated audio. The face- and speech-based opponent was accompanied by a semi-realistic animated human face. The computer opponents were programmed to cooperate in four out of six trials. While the face- and speech-based computer opponents were largely unable to garner trust in players (likely stemming from an uncanny valley effect), the text-based computer opponent was able to encourage the same rates of cooperation in subjects as human opponents. This finding suggests that human players are able to respond prosocially to some forms of computer opponents.

The Ultimatum Game is similar to the Prisoner's Dilemma in that they both explore players' intentions to accept or reject the formation of a social contract. However in the Ultimatum Game, two players must decide how to divide a sum of money between each other. In an experiment conducted by Rilling, Sanfey, Aronson, Nystrom, and Cohen (2004), both the Prisoner's Dilemma and the Ultimatum Game were used in order to gain insight into the difference between interactions with a human or computer partner in terms of "theory of mind," or one's conception of another person's thoughts and mental state in a social capacity. In this version of the Prisoner's Dilemma, cooperative payoffs were inflated to encourage cooperation. Results indicated that subjects are more likely to accept unfair behavior from a computer player rather than a human player (Rilling et al., 2004). This suggests that human subjects do not hold the computer to the same social constructs they hold other humans to, alluding to the issue of not considering the computer used in this experiment as a socially equal opponent. Similar to the Ultimatum Game, the Trust Game leaves two players the task of splitting a resource, with one player ultimately deciding how much each player receives (Anderhub, Engelmann, & Güth, 2002). In McCabe, Houser, Ryan, Smith, and Trouard

(2001), subjects played the Trust Game in both a human and a computer player condition. The computer player used a probabilistic model, the choice probabilities of which were shown to the subjects. Functional MRI (fMRI) results uncovered neural correlates indicating that the active brain areas involved varied between the two opponent types. While both opponents engage the prefrontal cortex in order to form a mental picture of the other player's state of mind, human opponents evoked higher prefrontal cortex activation and more cooperation attempts in some subjects.

It is important to note that these example experiments using the Prisoner's Dilemma and the Ultimatum Game paradigms have utilized either set strategies or preprogrammed responses in their computer agents. However, an agent with an adaptive strategy, one that learns in real-time while playing a game with another, might produce results that not only engage the human player in a higher capacity, but may also emulate human players enough to evoke strong social responses that influence behavior during play. Along these lines, Asher and colleagues introduced embodied, neurobiologically-plausible models of action selection and neuromodulation with the ability to adapt to their opponent's behavior while playing the game Hawk-Dove (Asher, Zaldivar, Barton, Brewer, & Krichmar, 2011; Asher, Zhang, Zaldivar, Lee, & Krichmar, 2012). These models incorporated the roles of the dopaminergic and serotonergic neuromodulatory systems in tracking expected rewards and costs, respectively. Because of their adaptive nature and physical embodiment, these models evoked interesting, strong, and complex responses from subjects. The Hawk-Dove game consisted of a human and a neural agent choosing to either share (Display) or fight (Escalate) for a valued resource. Whereas an unchallenged escalation (one subject escalates, the other displays) resulted in the escalating subject receiving the total value of the resource, a challenged escalation (where both subjects escalate) resulted in a costly penalty. If both subjects displayed, they shared the value of the resource. Thus, this paradigm optimizes investigation into risk-taking and cooperative behavior.

In order to study the effects of embodiment, subjects played Hawk-Dove games against both a simulated computer agent and an autonomous, physical robot (Asher, Barton, et al., 2012; Asher, Zhang, et al., 2012). In both cases, in order to probe the neuromodulatory mechanisms that give rise to cooperative and competitive behaviors, subjects played against a model with an intact serotonergic system and a lesioned serotonergic system, the latter of which typically made the agent play more aggressively. To impair the human player's serotonergic systems, subjects underwent an acute tryptophan depletion (ATD) procedure, which temporarily lowered serotonin levels and has been shown to reduce cooperation in the Prisoner's Dilemma game (Wood, Rilling, Sanfey, Bhagwagar, & Rogers, 2006). Subjects adjusted their strategies depending on the type of agent they played. Subjects exhibited a significant shift from a Win-Stay-Lose-Shift (WSLS) strategy against an intact agent to a Tit-for-Tat strategy against an agent that was more aggressive due to lesions of its simulated serotonergic system. This strategy change suggested that subjects were sending a message to the aggressive agent that they were being treated unfairly.

In the Asher et al. study, two groups best described individual subject's responses. ATD caused some subjects to be more aggressive, but others to be less aggressive, as seen by their probability to escalate a fight. A similar trend of two polarized subject groups was observed when considering the effect of physical embodiment on game play. This study showed that an adaptive agent could evoke strong, varied responses in subjects

(Asher, Zhang, et al., 2012). This suggests that there might be underlying biological or experiential factors leading to subject tendencies and or phenotypes in social situations.

The Stag Hunt has recently been used to test theory of mind assumptions, both through modeling and by human subjects against computer agents (W. Yoshida, Seymour, Friston, & Dolan, 2010; Wako Yoshida, Dolan, & Friston, 2008). In Yoshida and colleagues' experiment, subjects played Stag Hunt with a computer agent possessing one of three levels of sophistication, defined by the number of levels of reciprocal belief inference used by the model. Players were not aware of the level of sophistication used by the agent. Their fMRI results showed that rostral medial prefrontal cortex, a brain region consistently identified in psychological tasks requiring mentalizing, had a specific role in encoding the uncertainty of the other's strategy, and that the dorsolateral prefrontal cortex encoded the depth of recursion of the strategy being used. Their study demonstrates that socioeconomic games like the Stag Hunt and sophisticated computer agents can provide an excellent environment for probing social contracts, decision-making, and theory of mind.

Theory of Mind and Autism Spectrum

Autism spectrum disorders (ASD) are typically characterized by difficulties in communicating and forming relationships with other people (Wang & Doering, 2015). Beginning with the shift away from discrete diagnoses to the umbrella term featured in the DSM-V, the definition of autism has been formally changing to encompass a wider variety of disorders of social impairment including Asperger's, childhood disintegrative disorder, and pervasive development disorder ("Autism Spectrum Disorder," 2013). The DSM-V characterizes autistic impairment as a deficit in social communication and interaction or repetitive or restrictive behavioral patterns that first appear early in the developmental period, cause impairment in typical functioning, and cannot otherwise be explained by intellectual disability (American Psychiatric Association, 2013). Typically, ASD is diagnosed using behavioral assessments in the first two to four years of life (Lord, Rutter, & Le Couteur, 1994).

Social deficits in individuals with ASD are in part attributed to impaired theory of mind (ToM), the ability to infer what another person is thinking, feeling, or perceiving (Simon Baron-Cohen, Leslie, & Frith, 1985; Boucher, 2012; Chevallier et al., 2014; Kana, Keller, Cherkassky, Minshew, & Just, 2009; Mason, Williams, Kana, Minshew, & Just, 2008; Senju, 2012; Spek, Scholte, & Berckelaer-Onnes, 2010; Zalla, Miele, Leboyer, & Metcalfe, 2015). Common assessments of ToM abilities include tasks in which individuals must read stories and infer mental states of the characters (e.g. the Faux-Pas test and the Strange Stories test; (Spek et al., 2010; S. White, Hill, Happé, & Frith, 2009). While these tasks are successful in revealing positive correlations between utilization of ToM and the extent of autistic affectedness, narrative-based approaches rely on verbal ability and are abstracted by fiction. The field would benefit from the introduction of dynamic tasks that can be conducted non-verbally, engaging individuals with potentially limited verbal skills in social situations while minimizing direct social contact, in order to probe critical thinking and ToM in real-time.

Theory of mind (ToM) is the ability to imagine the beliefs, intentions, and mental states of social partners (Leslie, Friedman, & German, 2004). Investigation of a ToM network in the brain is an enduring topic of research that is focused on understanding the neural correlates of interpersonal interactions that are an inherent part of day-to-day life.

Research of the ToM network is crucial in gaining insight into the causes of and treatments for social impairments such as ASD (Simon Baron-Cohen et al., 1985; Stone, Baron-Cohen, & Knight, 2013).

ToM occurs in both intentional and subconscious forms, referred to as "explicit" and "implicit," respectively (Frith & Frith, 2008; Schuwerk, Vuori, & Sodian, 2014). Most commonly, ToM is probed through narrative tasks that *explicitly* require the participant to imagine what another person is thinking. One common example, the False Belief task, tests a subject's ability to represent another person's knowledge separately from their own, allowing for the fact that other people may not know what you know. The False Belief task is trivial for typical subjects, yet young children and some individuals with ToM impairments are not able to pass it (Apperly et al., 2004; Helming et al., 2014; S. White et al., 2009). While less commonly utilized, ToM can also be investigated through non-narrative tasks that provide their context *implicitly* through an animation or a game environment. Through clips of animated shapes using biological or random motion, Castelli et al. was able to probe ToM in both neurotypical (NT) and cognitively able subjects with autism, identifying differences in brain activity within the ToM network (Castelli, 2002). Similarly, Yoshida et al. utilized the game theoretic Stag Hunt task, evoking ToM implicitly in the preferred cooperative outcome (Yoshida, Seymour, et al., 2010). The addition of eye-tracking hardware to ToM studies can also be a useful probe into cues that signal ToM, especially in studies that incorporate individuals with autism because attempts to engage in ToM can be interpreted through eye gaze. For instance, individuals with autism are more likely to look at the mouth and less likely to look at the eyes of a social partner, indicating that the cues they are using are different from neurotypical individuals (Boraston & Blakemore, 2007). Individuals with autism can often successfully describe other individuals' mental states when explicitly asked, but will fail to use ToM inferences when performing a task that benefits from implicit ToM (Schuwerk et al., 2014).

As a result of the abstract nature and subsequently tenuous understanding of ToM, it is often a useful tactic to explore and reinforce hypotheses through computational modeling of its functional or physical components. Because ToM is applicable in social situations of all scales, it is useful to study both groups of agents interacting in a simulation such as Pynadath and Marcella's PsychSim model of ToM agents (Pynadath & Marsella, 2005), which models personality traits and behaviors of a group of people in order to simulate bullying in schools, as well as the individual's ToM experience in order to match collected behavioral data (Bosse, Memon, & Treur, 2007; El Kaliouby & Robinson, 2005; Friedlander & Franklin, 2008; Si, Marsella, & Pynadath, 2010; A L Thomaz, Berlin, & Breazeal, 2005; Andrea Lockerd Thomaz, Berlin, & Breazeal, 2005).

Amongst the most promising of these studies are those that are neurobiologically inspired and explore ToM down to the neurochemical processes involved. For instance, Abu-Akel and Shamay-Tsoory's neural network model of ToM that attempts to model specific brain areas incorporated in the mentalizing network (Abu-Akel & Shamay-Tsoory, 2011), as models of this type align themselves most closely with the machinery upon which they are based. In a specific example of ToM modeling work related to the present study, Hampton et al. conducted a study utilizing modeling and game theory to probe mentalizing behavior in an fMRI setting with the intention of matching up the model's activity with neural data (Hampton, Bossaerts, & O'Doherty, 2008). Subjects played an economic game in the scanner against other human players, and the data collected from this experiment was shown to match the data of the model playing the game. While this study was successful in attributing decision-making activity to mPFC, bilateral STS, and various other areas, the authors only used human players during the experiment. While they did not use their model as a player, they expressed the desire to see studies in the future that utilized adaptive agents in this capacity, a primary goal of the present work.

Recent research has indicated that differing levels of autistic traits can be found throughout the general population (Simon Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Ruzich et al., 2015), indicating that the spectrum of ASD may extend beyond clinical diagnoses and into subclinical populations. In an effort to study varying levels of autistic traits, the Autism Quotient (AQ) test, a 50 question Likert-based survey querying the presence or absence of various behaviors typically associated with ASD was created (Simon Baron-Cohen et al., 2001). The AQ test allows identification of the presence of autistic traits, such as ToM impairment, in the general population and can be used to bridge the gap between subclinical and clinical populations.

In research closely related to the AQ test, the Empathizing/Systemizing Quotient tests are also used to assess the degree of autistic traits in general population individuals (S. Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003; Simon Baron-Cohen, 2009; Focquaert & Vanneste, 2015; Wheelwright et al., 2006). Individuals with autism and high levels of autistic traits tend to be described as “high systemizing,” meaning that such individuals tend to analyze systems by rules and patterns, whereas individuals with low levels of autistic traits tend to be “high empathizing,” meaning that they tend to predict and respond to mental states and emotions (Focquaert &

Vanneste, 2015).

Neural basis of ToM and Autistic Impairment

Research on the neural expression of the ToM network is crucial in gaining insight into the causes of and treatments for social impairments such as Autism Spectrum Disorder (ASD) (Baron-Cohen et al., 1985; Stone et al., 2013), which is often accompanied by impairments in ToM ability. As ToM is a high-level phenomenon that exhibits distinct characteristics through both behavior and neural activity, identifying the neural correlates, including both specialized areas of the brain and their corresponding networks, is key in understanding how ToM is represented and carried out in each specific situation. As a result of the abstract nature and subsequently tenuous understanding of ToM, it is often a useful tactic to explore and reinforce hypotheses through 1) noninvasive brain imaging measures and 2) computational modeling of its functional or physical components. There is a core network of brain areas that plays a role in ToM tasks and is therefore categorized as ToM areas: TPJ (R. Saxe & Powell, 2006), APC (Gallagher & Frith, 2003), STS (Allison, Puce, & McCarthy, 2000; Beauchamp, 2015; Deen & Saxe, 2012; von dem Hagen et al., 2011), mPFC (Yoshida et al., 2010), and ToM supplementary areas such as the amygdala (Fine, 2001) and OFC (Gallagher & Frith, 2003).

In previous work, mPFC has been found to be engaged in perspective taking (Hari & Kujala, 2009), shared attention to a goal (Rebecca Saxe, 2006), or as an integration area for social information (Van Overwalle, 2009). TPJ has traditionally been viewed as an area specialized for ToM, involved in mentalizing or imagining the goals and desires of a social partner (Hari & Kujala, 2009), perhaps in an effort to understand the mental states (Rebecca Saxe, 2006) or intentions of others (Van Overwalle, 2009). The APC and STS are two additional areas that have been implicated in attempting to understand the intentions of others (Gallagher & Frith, 2003; Hari & Kujala, 2009), especially utilizing biological motion cues in the case of STS (Allison et al., 2000; Rebecca Saxe, 2006; Van Overwalle, 2009). The bilateral temporal poles have shown activity in ToM experiments, particularly narrative driven (Olson, Plotzker, & Ezzyat, 2007), in regards to attributing mental states (Hari & Kujala, 2009) and relating to other people, perhaps in ways informed by one's personal memories of similar experiences (Gallagher & Frith, 2003). OFC and amygdala have both shown to be active in emotional contexts (Camille, 2004; Coricelli et al., 2005; Moll, de Oliveira-Souza, Bramati, & Grafman, 2002), emotion having a strong influence on ToM and social interaction in general, a compelling argument as to why these areas are relevant in this context. OFC activity correlates with the influence of moral responsibility on emotion (Moll et al., 2002) (e.g. guilt), while the amygdala appears to be active in situations involving the effect of one's emotions on social judgments (Hari & Kujala, 2009).

As mentioned above, impairment of ToM in varying degrees of severity is a common deficit found in people with ASD (Matthews et al., 2012). As autism is characterized by a difficulty interacting with other people and understanding their mental states (Eigsti & Shapiro, 2003), the primary ToM network and its interaction with other areas of the brain is commonly discussed alongside autism (Simon Baron-Cohen et al., 1985). While ToM deficits in individuals with ASD are pervasive early in childhood and adolescence, through experience and therapy, it is not uncommon to see improvements in autistic individuals such that they exhibit a definite improvement in the ability to cope with social situations by late adolescence and adulthood (White et al., 2014). However, it has recently been shown that while these deficits are not as apparent behaviorally in advanced years, the neural correlates of these deficits and their subsequent compensations are still observable using fMRI (S. J. White et al., 2014). White et al. identified an hyperactivation of the mentalizing network, characterized by the regions of interest (ROIs) mPFC, posterior cingulate cortex (PCC), TPJ, and temporal poles in an explicit, narrative-driven fMRI task of ToM in comparison to NT subjects (S. J. White et al., 2014), in agreement with the results of other similar tasks (Mason et al., 2008), and a complementary pattern of hypoactivation has been shown in individuals participating in implicit ToM tasks (Castelli, 2002).

The field of autism and ToM could benefit largely from further investigation into this supported theory of ToM network activation as a function of explicit vs. implicit tasks. White et al. hypothesize that overactivity of the neural correlates of ToM in people with autism may be correlated to explicit tasks because they are made aware of the task demands and perform strong compensation for deficits, whereas implicit tasks may not be sufficient to cue a ToM response in subjects with ASD, leading to decreased activity in ToM areas of the brain; both hypotheses have been supported by findings from related studies (Castelli, 2002; Kana, Keller, Minshew, & Just, 2007; Koster-Hale, Saxe, Dungan, & Young, 2013; M. V. Lombardo et al., 2010; Mason et al., 2008; S. J. White et al., 2014). Coupled with the previously discussed concept that the general population expresses varying levels of autistic traits, the investigation of brain activity of adults in the general population during ToM tasks is a promising venture capable of providing insight into the neural correlates of ToM and autism by means of comparison to differential activation patterns.

CHAPTER 1: Social Contracts and Human-Computer Interaction with Simulated Adapting Agents

In the present study, we are interested in moving beyond games that focus on the competition between players, to explore teamwork and social signaling among players by using the socioeconomic game known as the Stag Hunt (Skyrms, 2004). In the game of Stag Hunt, two players decide whether to hunt a high-payoff stag cooperatively or a low-payoff hare individually. As described in detail in Scholz and Whiteman (2010), the risk in this game is that both players must decide to hunt the stag in order to catch it. In the case that both players hunt the stag, both are awarded a high payoff. However, if only one player decides to hunt the stag while the other hunts a hare, the player who hunted the stag gets no payoff and the player who hunted the hare obtains a small payoff. Thus, success in the Stag Hunt requires the ability to make a social contract with another player and form a representation of another's intentions. A description of this work can be found in Adaptive Behavior (A B Craig, Asher, Oros, Brewer, & Krichmar, 2013)

A major goal of the study presented in this chapter is to show that an agent with the ability to adapt to another player's gameplay more effectively challenges a subject. In many Human Computer Interface (HCI) games, subjects play against computer opponents with static strategies, which may not challenge subjects in a natural way. A simulated agent with the ability to adapt to its opponent's behavior has the potential to evoke more complex and interesting results in subjects than these set-strategy agents used in the studies described above. Such an adaptive system may be a more informative probe for investigating human behavior under challenging conditions. The use of adaptive agents provides a controlled way to make subjects believe they are playing against an intelligent opponent. Moreover, incorporating the adaptive behavior observed in subjects into future simulated agents may lead to HCI systems that interact more naturally with people.

To move beyond the more simplistic and commonly used paradigm of game play against agents with set-strategies, the present study investigated the social and behavioral effects of an *adaptive* agent on human subjects within the highly social Stag Hunt game environment. In order to compare pre-set and adaptive agent paradigms, human subjects played a computerized version of the game with five different strategies: exclusive hare hunting, exclusive stag hunting, random hunting, Win-Stay, Lose-Shift (WSLS) hunting, and an adaptive agent. The adaptive agent was implemented with an Actor-Critic model that took into account the costs and benefits of moves. Our results show that such an adaptive agent is able to evoke a response in subjects that is significantly different from those produced by set-strategy paradigms. Subjects spend more time and effort when playing against an adaptive agent, following more complex paths to their targets. Thus, such adaptive agents have the potential to be used in social situations as a partner or opponent akin to another human player, while allowing for greater control.

1.1. Methods

1.1.1. Human Participants.

Forty subjects (age range: 18-25) were recruited through an online database maintained by the Experimental Social Science Laboratory (ESSL) at the University of California, Irvine (UCI). The subject database is comprised of currently enrolled undergraduate and graduate students from UCI who have volunteered to be contacted for and participate in socioeconomic experiments within the UCI School of Social Sciences. In this recruiting database, there is no screening for race, gender or other background characteristics. Subjects had not previously participated in the same experiment. The experimental protocol was approved by the Institutional Review Board at University of California, Irvine, and informed consent was obtained from all subjects. Two subjects did not appear to understand the instructions for the majority of the experiment; their data was removed before analysis.

1.1.2. Computer Interface for the Stag Hunt

Subjects played a variant of the Stag Hunt game against simulated agents, which was similar to the game used by Yoshida and colleagues (W. Yoshida, Seymour, et al., 2010). This version of the Stag Hunt game differed from the traditional version by adding a spatial and temporal component to the game. The spatial component consisted of a game board with tokens, both for the players and for the stag and hare prey, such that the players needed to traverse squares on the board in order to reach and capture their prey. The temporal component was a byproduct of this game environment in that it took a variable amount of time in each game to reach and capture prey. This non-standard approach was used in order to provide more measurable differences in human behavior beyond the record of the action choices themselves (e.g. reaction time, number of turns, path on gameboard, etc.). However, the present version retains the stag and hare equilibriums of the original version of Stag Hunt.

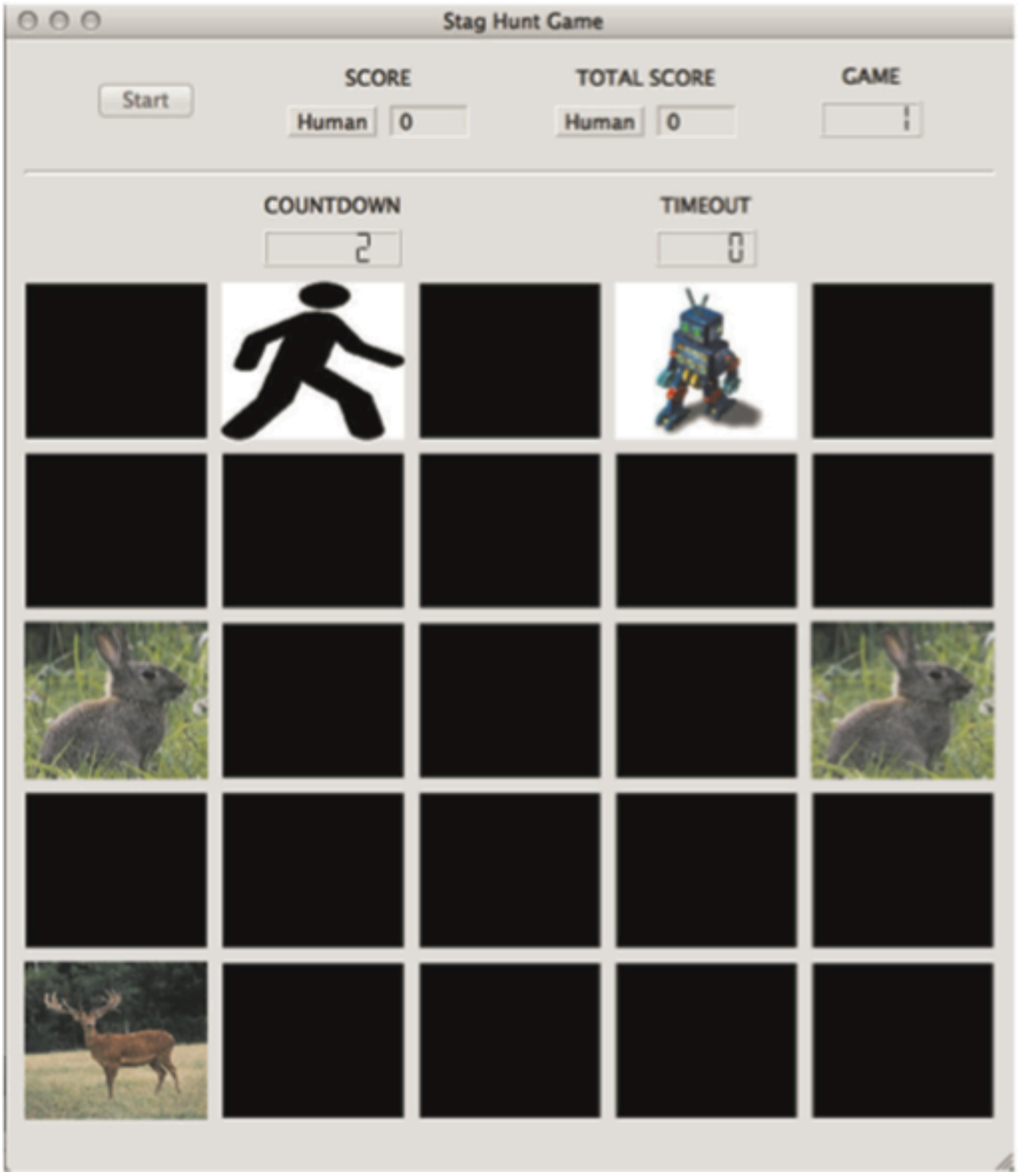


Figure 1.1. Screenshot of Stag Hunt game board.
The game board included a 5 × 5 grid of spaces upon which the player (stick figure image), agent (robot image), stag (stag image), and hare (hare image) tokens resided. The screen included a button to start the experiment, the subject's score for the round, the subject's overall score for the experiment, the game number within the round, a 3-second countdown to the start of the game, and a 10-second counter monitoring the game's timeout. At the beginning of each game, the locations for the stag, player, and agent tokens were randomly placed along either the top row, bottom row, or middle column at least one square away from each other. The initial positions of the hares were fixed in the locations shown above for all games. The player and agent could move one square at a time towards their goal at the start of the game, while the targets remain fixed.

The computer interface consisted of a 5x5 grid on which a stag token, two hare tokens, a subject token, and an agent token were placed (Figure 1.1). The two hare tokens were placed on the middle square of the left and right columns for every game while the stag, subject, and agent tokens were randomly placed on a square residing within the first row, last row, or center column of the game grid. This precaution ensured that the players and the stag would not begin a game right next to a hare. Player tokens were prevented from being initially placed directly next to or on top of a stag or each other.

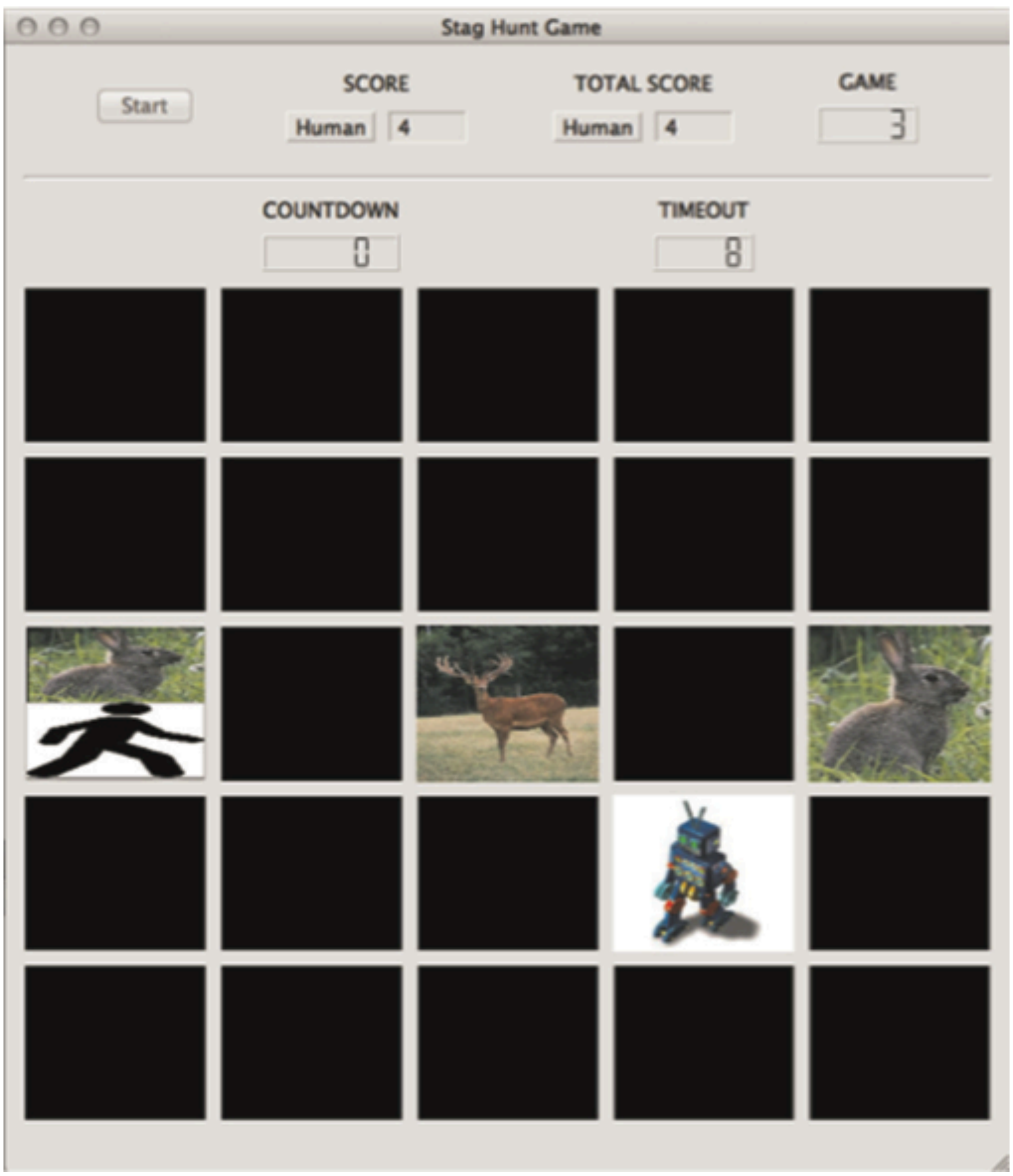


Figure 1.2. Screenshot of hare capture.
Players moved towards a target by performing consecutive left mouse clicks on adjacent squares until they had arrived at their target. In order to catch a hare, the player needed to be on top of the hare so that the image displayed both the player's and the hare's tokens. The player then performed a right mouse click on top of the current square to catch it. In the case that both players were on a hare square, the first player to click on the hare caught it. When a player caught a hare, that player won one point and the current game ended.

Each participant controlled the subject token through left mouse clicks to adjacent squares on the grid to hunt either the stag or hare token. Moves were executed simultaneously between players (i.e. were not limited to turns), and the subject's moves took effect instantaneously. Computer agents moved every 600 ms, which was shown in software testing to create a reasonable level of difficulty (assessed by near-equal agent/subject point totals in non-expert players). Subjects were capable of moving quickly (~200 ms), but often took more time in deciding moves. In order to hunt a hare, the subject token needed to occupy the same square as a hare token (Figure 1.2). A subject made a right mouse click on the currently occupied hare square to catch the hare. In the event that both players tried to catch a hare at the same time, the player that made the first click caught the hare. In order to hunt a stag, the subject and agent tokens needed to occupy squares adjacent to the stag token vertically, horizontally, or diagonally (Figure 1.3). A subject made a right mouse click on the stag square in order to catch it. It was not sufficient for both players to merely be next to the stag; they both needed to indicate their intentions to catch the stag. As soon as a hare or stag was caught, the game ended. Catching a stag awarded each player four points, while catching a hare awarded the successful player one point and the unsuccessful player zero points (see payoff matrix in Table 1.1).

Table 1.1. Payoff matrix of Stag Hunt.

	Agent hunts Stag	Agent hunts Hare
Player hunts Stag	Agent: 4 pts Player: 4 pts	Agent: 1 pt Player: 0 pts
Player hunts Hare	Agent: 0 pts Player: 1 pt	First to catch: 1 pt Other: 0 pts

During the games, the subjects saw their total scores for the current round as well as a 10 second countdown timer for each game that provided a time limit for each game. If a game lasted over 10 seconds, no payoffs were given. At the end of each round, the subjects were shown their total scores summed over all rounds already played. Subjects were not shown the score of the agent in order to prevent unnecessary competition.

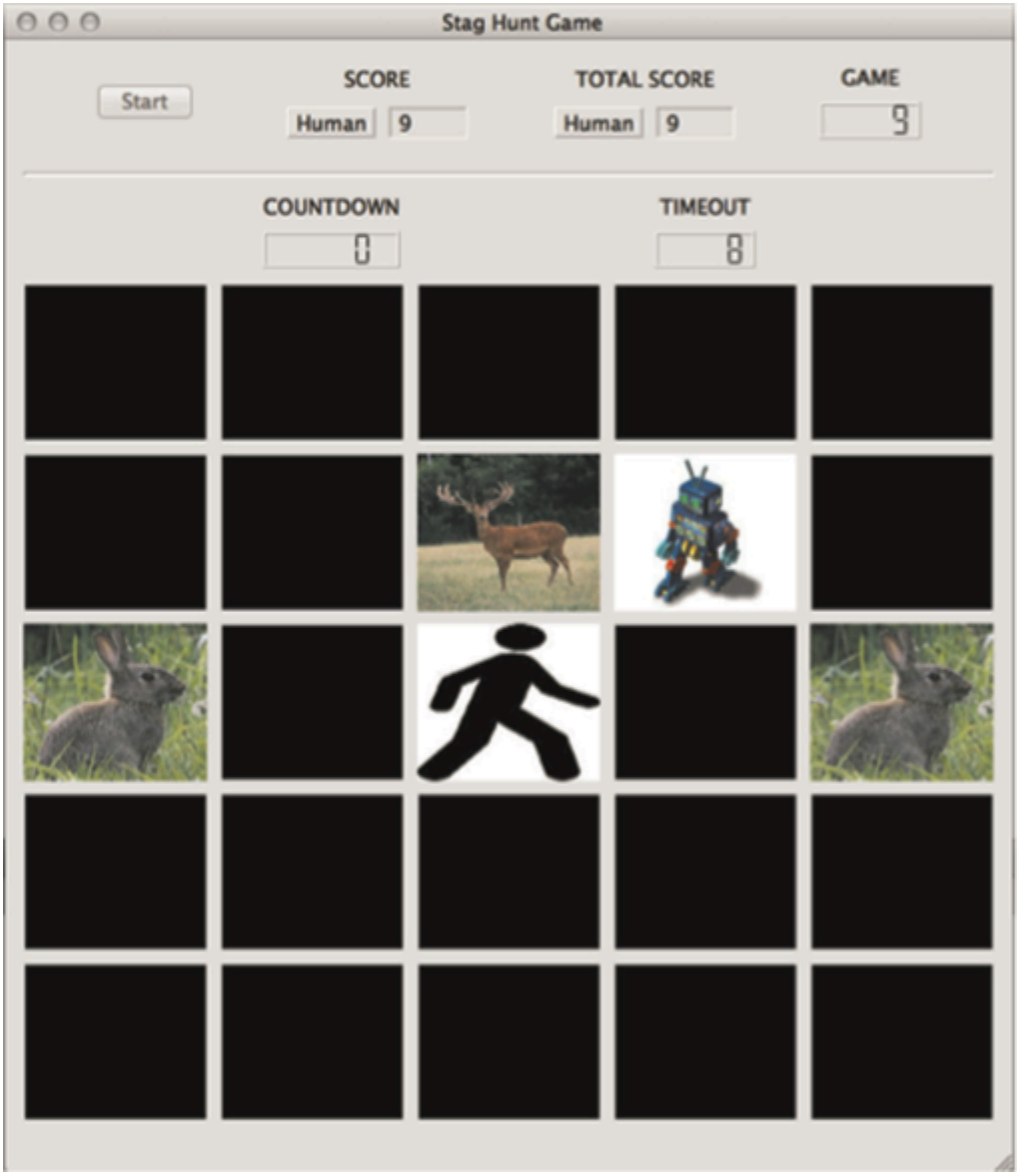


Figure 1.3. Screenshot of stag capture.
In order to catch a stag, both the player and agent tokens needed to be in squares adjacent to the stag token, whether horizontally, vertically, or diagonally adjacent. Both the player and the agent required the intention of catching a stag. It was not sufficient to simply pass next to the stag while the other player intended to catch it. Once both players were adjacent to the stag and had the intention to catch the stag, the human player performed a right mouse click on top of the stag in order to catch it. Catching a stag awarded both players four points each.

1.1.3. Agents for the Stag Hunt

For each of the 250 games of the Stag Hunt, the agent played one of the following five strategies: EQStag, EQHare, Random, Win-Stay-Lose-Shift (WSLS), and Adapt. EQStag agents always hunted stags, while EQHare agents always hunted hares. The Random agent had an equal probability of hunting a hare or a stag in each game. The WSLS agent chose either hare or stag randomly in its first game, switching to the other target after losing a game and repeating its choice after winning a game. The Adapt agent began its first game with no choice preference or strategy, and developed its strategy through an Actor-Critic model that will be described below. The rounds were presented in random order for each subject, and all subjects played against every agent strategy. No significant order effects were found.

During the round in which the subject played against the Adapt agent, an Actor-Critic model was employed, which learned the appropriate actions based on the rewards and penalties acquired during a series of Stag Hunt games.

The model updated state tables for a Reward Critic, Cost Critic, and Actor. Each state was designated by: 1) the player's distance from hare, 2) the agent's distance from hare, 3) the player's distance from stag, and 4) the agent's distance from stag. The distances were calculated using Euclidean distance and then truncated to the nearest integer value. Player tokens could be, at most, five squares from the stag and three from the nearest hare, so there were 225 possible states in each table.

The Reward Critic state table contained a weight that corresponded to the expected reward at the current state. Reward was defined as the payoff received at the end of a game as given by the payoff matrix (see Table 1.1). Similarly, the Cost Critic state table contained a weight that corresponded to the expected cost at the current state. Cost was defined as the perceived loss on a hunt. For example, if the Agent was hunting a stag and the human caught a hare, the cost would be -4 (see Table 1.1). The Actor state table contained two weights for each state: one for the likelihood to hunt hare and the other for the likelihood to hunt stag in a given state. The Adapt agent was naïve for each subject at the beginning of the experiment, meaning that the state tables were initialized to zero.

After each move made by either player, the Actor-Critic model state tables were governed by the following equations.

The Actor-Critic weights depended on a delta rule that calculated an error prediction:

$$\delta(t) = r(t) + V(s,t) - V(s,t-1)) \tag{1}$$

where $r(t)$ was either the reward or cost at time t, $V(s,t)$ was the Critic's weight at state s, at time t, and $V(s,t-1)$ was the Critic's weight for the previous timestep. $r(t)$ for the Reward Critic was given as:

$$r_{rwd}(t) = \begin{cases} 4;\ if\ caught\ stag\ at\ time\ t \\ 1; if\ caught\ hare\ at\ time\ t \\ 0; otherwise \end{cases} \tag{2}$$

$r(t)$ for the Cost Critic was given as:

$$r_{cost}(t) = \begin{cases} -4;\ if\ hunting\ stag\ and\ other\ player\ caught\ prey\ at\ time\ t \\ -1;\ if\ hunting\ hare\ and\ other\ player\ caught\ prey\ at\ time\ t \\ 0; otherwise \end{cases} \tag{3}$$

The Critic's state table was updated by:

$$V(s, t+1) = V(s, t) + \delta(t) \tag{4}$$

Equations 1-4 were applied after each move to update the weights in the Reward and Cost Critic state tables.

The Actor's weights were updated according to Equations 5 and 6. Equation 5 is given for the condition in which the Adapt agent hunted a hare.

$$\begin{aligned} V(hare, s, t+1)) &= V(hare, s, t) + 1 - p[hare] * \delta(t) \\ V(stag, s, t+1)) &= V(stag, s, t) - p[stag] * \delta(t) \end{aligned} \tag{5}$$

$V(hare, s, t)$ was the Actor's state table value for hunting a hare in state s at time t. Likewise, $V(stag, s, t)$ was the Actor's state table value for hunting a stag in state s at time t. $\delta(t)$ was the delta value from both the Reward and Cost Critics. Equation 6 is given for the condition in which the Adapt agent hunted a stag.

$$\begin{aligned} V(stag, s, t+1)) &= V(stag, s, t) + 1 - p[stag] * \delta(t) \\ V(hare, s, t+1)) &= V(hare, s, t) - p[hare] * \delta(t) \end{aligned} \tag{6}$$

Equations 5 and 6 were applied for both the Reward and Cost Critic. The probability for hunting a hare, $p[hare]$, or stag, $p[stag]$, was given by the SoftMax function:

$$p[hare] = \frac{e^{V(hare,s,t)}}{e^{V(hare,s,t)} + e^{V(stag,s,t)}}$$

$$p[stag] = 1 - p[hare] \tag{7}$$

At each turn, Equation 7 was used to choose the agent's prey. The agent would then move one square closer to the stag, if stag was chosen, or one square closer to the nearest hare, if a hare was chosen.

1.1.4. Experimental Design

Data for each subject were collected simultaneously on forty Dell desktop computers in the ESSL, with each subject separated by privacy boards to prevent distraction and discussion between subjects. The subjects first watched a narrated PowerPoint presentation, which provided a standardized explanation of the purpose and instructions for the experiment. Subjects were informed at this time that they would receive both a baseline compensation for participation as well as an incentive payment that was dependent on their performance in the experiment game play. The subjects next participated in a training session in which they played ten games of the Stag Hunt against a random-acting agent; the results from these ten games did not count towards the subjects' final scores. The experimental session then consisted of 250 games of Stag Hunt, divided equally between five rounds. Each subject played the Stag Hunt game against all five of the computer agents (as discussed above) in rounds of 50 games, one round per agent type, with the rounds presented in a random order for each subject. Subjects were aware of switches between the agents, but they were not given any information on the agent's strategy. Data for each subject were saved to text files, which were then compiled using Netsupport School computer software.

Following completion of the experiment, all subjects received a US$7 standard payment for experimental participation as well as compensation based on their performance at the rate of US$0.02 for each point won during the experimental session. Four points were awarded for catching a stag, or US$0.08, one point for catching a hare, or US$0.02, and zero points for not catching a target or allowing the ten-second timer to run out during a game. End of experiment payments ranged from US$10 to US$21.

1.2. Results

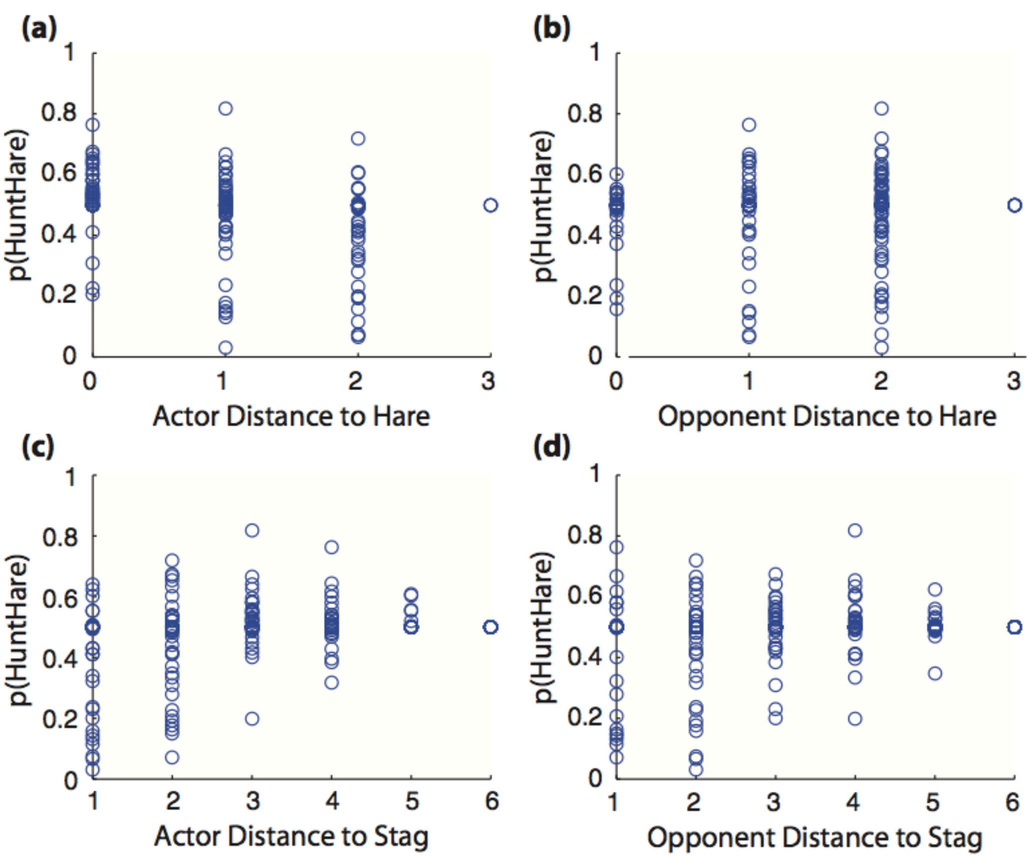


Figure 1.4. Scatter plot of Actor state table data.
Data for all subjects were taken from the Actor state tables of the Actor–Critic models used in the Stag Hunt experiment. (a) shows the probability for the adapting agent to hunt the closest

hare from each possible distance to closest hare target, while (b) shows the probability of the agent to hunt the closest hare from each possible distance of the subject to the closest hare target. (c) shows the probability for the agent to hunt the stag from each possible distance to the stag, while (d) shows the probability of the agent to hunt the stag from each possible distance from the subject to the stag.

The Adapt agent demonstrated the ability to adapt to the subjects' gameplay by taking into consideration the subjects' position with regard to game tokens. An analysis of the Actor state tables was performed to show the likelihood to hunt hare based on the distances of the Adapt agent and the subject from the stag and the closest hare. The Adapt agent was more likely to hunt a stag if it was further away from a hare (Figure 1.4(a)) or if the other player was further away from a hare (Figure 1.4(b)). Figures 1.4(c) and 1.4(d) show the Adapt agent was more likely to hunt a stag if either it or the other player were near a stag. These results show that the Actor-Critic algorithm was not only sensitive to its own position on the game board, but was also monitoring the other player's position.

Table 1.2. P-values for Wilcoxon rank-sum pairwise comparisons of average subject scores in each condition

	Adapt	EQHare	EQStag	Random	WSLS
Adapt		$< .0001$	$< .0001$	.9090	$< .0001$
EQHare			$< .0001$	$< .0001$	$< .0001$
EQStag				$< .0001$	$< .0001$
Random					$< .0001$

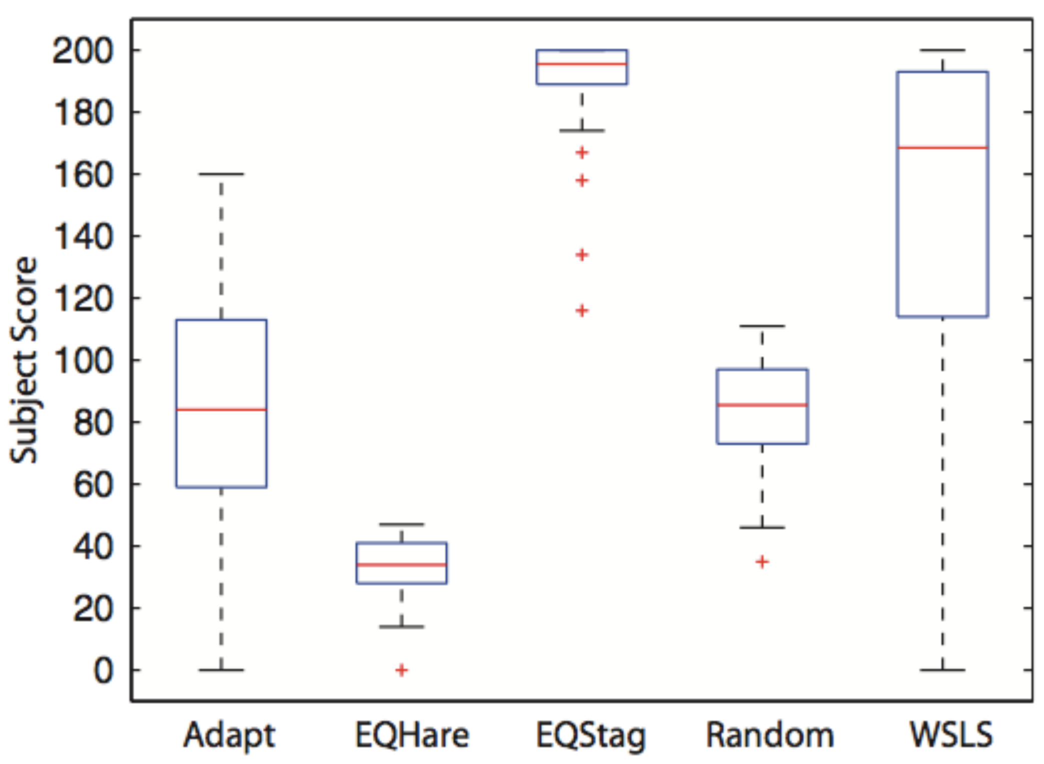


Figure 1.5. Subject scores against agent strategy.
For each boxplot, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually as ' + ' symbols. The data were not normally distributed; therefore subject performance against different agents was compared using Wilcoxon rank-sum tests (Bonferroni corrected, $p < .005$ was considered significant). The graph depicts the subject scores when playing with different agent strategies: Adapt, EQHare, EQStag, Random, and Win– Stay–Lose–Shift (WSLS). Scores were averaged over all subjects during the Stag Hunt experiment (Table 1.2).

Subject performance varied depending on the type of agent played (Figure 1.5; Table 1.2). In all cases, the agents' scores were similar to human scores, indicating that the two opponents were fairly matched. EQStag was shown to produce significantly higher overall subject scores than all other agent strategies, followed by WSLS, which produced significantly higher scores than the remaining three conditions. These high scores were due to the subjects gravitating towards cooperation and the high-payoff equilibrium of hunting stags. Subjects had the lowest scores against EQHare agents because they were forced to compete against their opponents for low-payoff hares.

Playing with Adapt and Random agents resulted in significantly higher scores than EQHare and lower scores than EQStag and WSLS, however they were not found to be significantly different from each other. Successfully hunting hare in all games would have resulted in a score of 50, while successfully hunting stag in all games would have resulted in a score of 200. Because subjects had scores higher than 50, yet lower than 200, this implies that subjects switched between hare and stag hunting against Adapt and Random agents rather than tending toward the hare or stag equilibrium. Furthermore, the wider range of scores when comparing Adapt to Random suggests that subjects had more difficulty figuring out the Adapt agent's strategy.

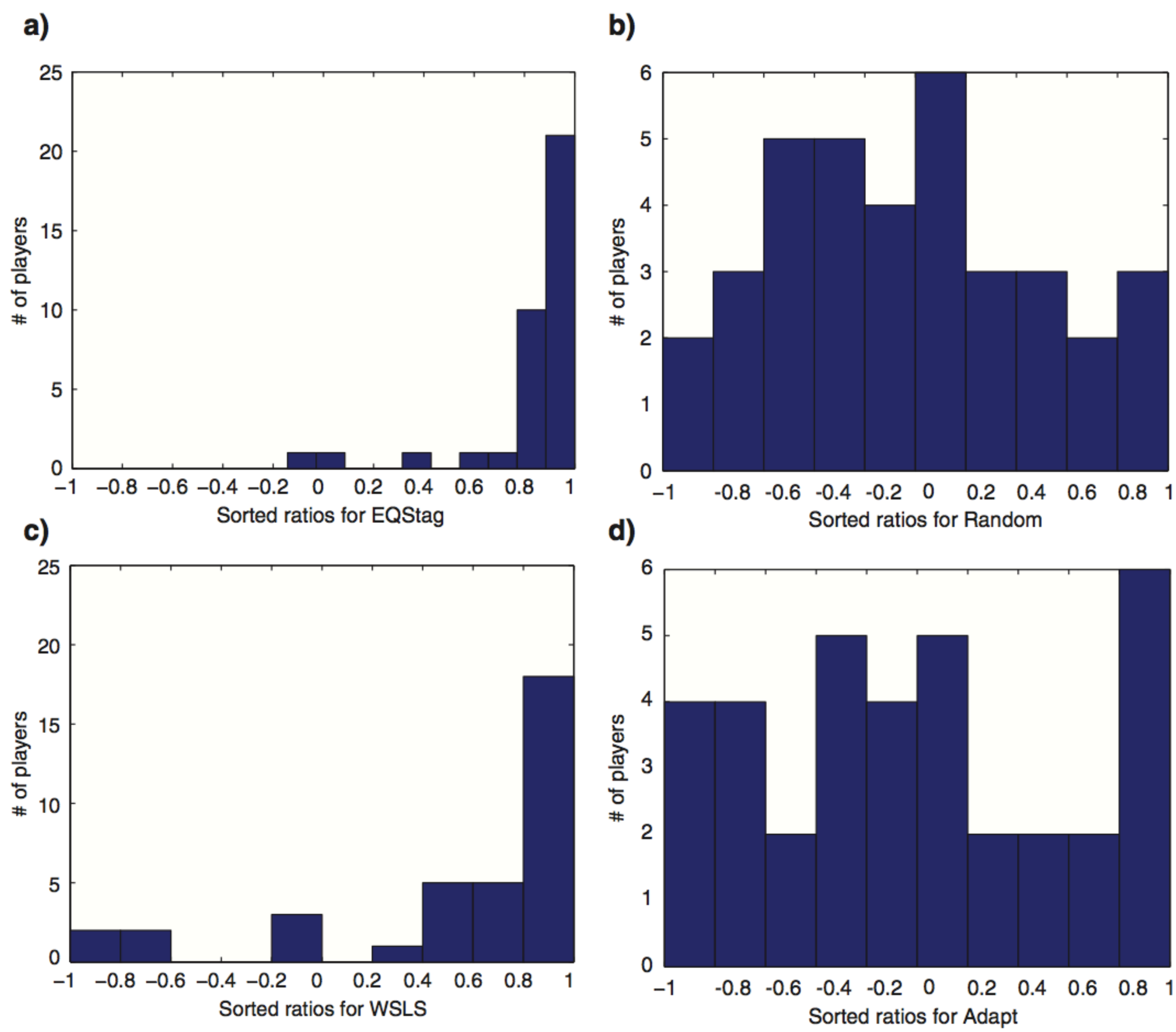


Figure 1.6. Ratio of stag to hare catches for all participants.
The ratio of stag to hare catches was calculated by the equation $q_{hare:stag} = (n_{stag} - n_{hare})/(n_{stag} + n_{hare})$, in which $q_{hare:stag}$ is the ratio for a given subject, n_{stag} is the number of stags captured during a given condition, and $-n_{hare}$ is the number of hares captured during a given condition. Values of positive one indicate exclusive stag hunting (EQStag), while values of negative one indicate exclusive hare hunting (EQHare). The histograms display the ratio data for (a) EQStag, (b) Random, (c) WSLS, and (d) Adapt agents. Note that the y-axis differs between Adapt/Random and EQStag/WSLS in order to better observe the shape of the data.

To understand how individual subjects altered their strategy when playing with different agents, we calculated the ratios of stag-to-hare catches for each subject in each condition (Figure 1.6). The ratio was calculated by using the equation,

$$q_{stag:hare} = (n_{stag} - n_{hare})/(n_{stag} + n_{hare}) \tag{8}$$

in which $q_{hare:stag}$ represents the normalized ratio of stags to hares, n_{stag} represents the total number of stags caught for that subject over all games in the condition, and n_{hare} represents the total number of hares caught for that subject over all games in the condition. Each ratio falls along a scale between negative one and positive one, negative one representing exclusive hare hunting and positive one representing exclusive stag hunting. In order to show the distribution of hunt behavior in subjects, Figure 1.6 shows how the subject hunted with an Adapt, EQStag, Random, or WSLS agent. We omitted the histogram for EQHare, as it was only possible for either player to catch a hare when playing with this strategy and thus all data points were at negative one. Also, two subjects were omitted from this analysis for not successfully catching any stags or hares in the Adapt and WSLS conditions. As expected, subjects showed a bias towards stag hunting when playing against EQStag (Figure 1.6(a)), which suggests that they found a high-payoff equilibrium. In Figure 1.6(b), subjects playing a Random agent had a somewhat normally distributed distribution of hunting tendencies with the peak being a mixture of stag and hare hunting. In Figure 1.6(c), subjects playing a WSLS agent had a bias towards stag hunting, as was also seen in EQStag and likely was a result of the high-payoff equilibrium. In Figure 1.6(d), subjects playing the Adapt agent had a trimodal distribution: 1) those preferring the cooperative equilibrium, 2) those preferring the non-cooperative equilibrium, and 3) those who were equally split between those two extremes.

Table 1.3. Color-coded chart of equilibrium alignment for individual subjects against each agent strategy.

The key shows the ratio, with green colors representing strong stag hunting equilibrium, and red colors representing strong hare hunting equilibrium.
Darker shades represent a stronger bias, and white represents minimal to no bias. The majority of subjects displayed positive/moderate ratios
throughout conditions, and those who displayed strongly negative ratios often remained negative or weakly biased throughout conditions.

Table 1.3 shows each individual subject's hunting bias for each condition, as indicated by their normalized ratios, with darker colors representing stronger biases toward stag or hare equilibrium. EQHare was omitted, because subjects could only capture a hare in this condition. As shown by the table, many subjects were biased to stag or hare hunting across different conditions. For example, subjects 11, 13, 15, 38, 55, and 57 remained strong stag hunters in multiple conditions, including the Adapt condition. Some subjects showed hare equilibrium tendencies across multiple conditions (i.e. Subjects 12, 14, 34, 41). These results suggest that subjects may have tendencies toward cooperation or non-cooperation. However, subjects 36, 37, and 52 tended toward hare hunting in the Adapt condition but not in other conditions, implying that the Adapt agent evoked a shift in strategy in some subjects.

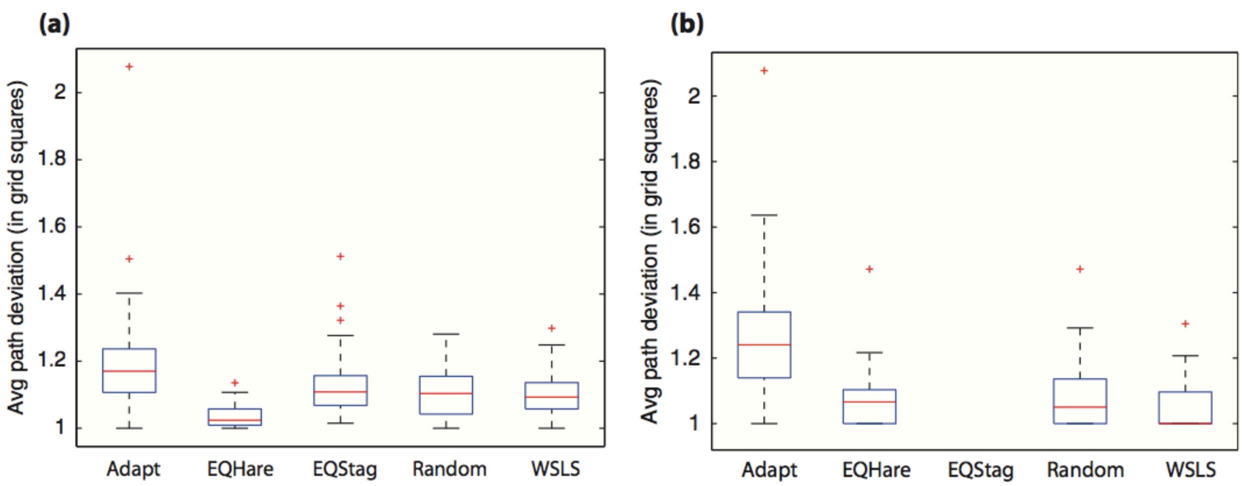


Figure 1.7. Average path deviation ratio over all subjects for each agent strategy. The boxplots have the same notation as in Figure 5. The data were not normally distributed; therefore subject performance against different agents was compared using Wilcoxon ranksum tests (Bonferroni corrected, $p < .005$ was considered significant). The length of the direct path to the target was calculated by measuring the distance between the first and last moves for each game of each subject. That distance was subtracted from the subject's total distance traveled in each game calculated by summing the distances between each move. Those differences were used in the above graphs as the average path deviation for each agent strategy: Adapt, EQHare, EQStag, Random, and Win–Stay–Lose–Shift (WSLS). (a) The average path deviations over all games and strategies. (b) The average path deviations for only the games in which the subject did not successfully catch a stag or hare, in other words losing the game (Table 1.5).

Subjects' paths were analyzed to determine the directness of their movements by measuring the amount of deviation from a direct path connecting their first movement toward their final destination at the end of the game, referred to as the "direct path". The games analyzed were for all outcomes (Figure 1.7(a); Table 1.3), and games in which the subject failed to catch either a stag or a hare (Figure 1.7(b); Table 1.4). Failures, in particular, were analyzed because the path deviation provided extra information as to why the subject failed to catch a target, for example indicating an attempt to observe the agent, attempting to hunt the stag while the agent hunted hare, etc. Path deviation was calculated by finding the length of the direct path (distance between the first and last moves for each game of each subject) and subtracting that number from the subject's total distance traveled in each game (calculated by summing the distances between each move).

Table 1.4. p-values for Wilcoxon rank-sum pairwise comparisons of average subject path deviation ratio in each condition over all games.

	Adapt	EQHare	EQStag	Random	WSLS
Adapt		$< .0001$	.0419	.0054	.0052
EQHare			$< .0001$	$< .0001$	$< .0001$
EQStag				.4030	.3689
Random					.9090

All comparisons were performed on the average path deviation ratio for each subject per agent strategy. In the rank sum analysis for path deviation over all games, EQHare showed smaller average deviations from all other conditions. Adapt showed nearly significantly larger deviations from Random and WSLS. No other comparisons were shown to be significantly different. However, in the rank sum analysis for losses (Figure 1.7(b); Table 1.4), Adapt was found to have significantly larger path deviations from all other agent strategies. These results might indicate that subjects realized that the adaptive agent's actions were malleable depending on subject behavior, and therefore subjects attempted to guide the agent or wait for the agent to change its target. EQHare and Random were both shown to have significantly larger path deviations than WSLS, however they were not significantly different from each other. EQStag had no analyzable loss data because the only way to lose a game in EQStag is to time out. All timeout data was removed before analysis due to excessive skewing.

The path deviation of the Adapt agent was analyzed in the same way as the human data. Performing the path deviation analysis on the Adapt agent showed that the average deviation per game over all subjects is 1.5 units ($\sigma = 0.4$) (compared to the human's average path deviation in Adapt (~1.2)). For reference, each of the other agent types had an average path deviation of 1 (direct path) due to their inability to switch targets mid-game. In addition to using path deviation to give insight into the player's intention, move data in the Adapt condition was analyzed to see which player arrived at the stag first. Subjects arrived first at the stag 36% ($\sigma = 16\%$) of the time stags were caught. The indirect paths and tendency to get to the stag first on many games, may suggest that subjects were trying to guide the Adapt agent's behavior.

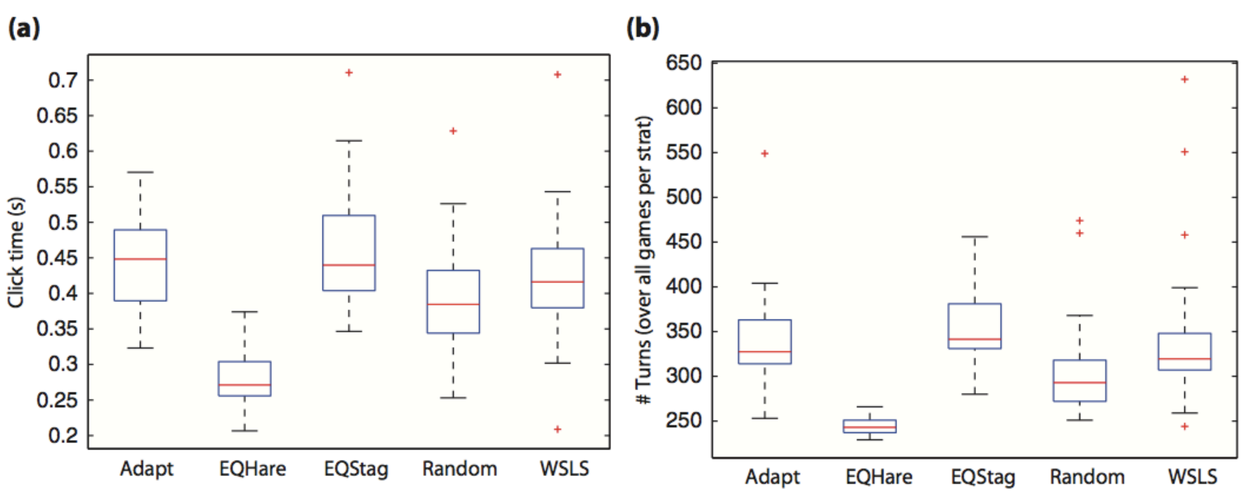


Figure 1.8. Average time between mouse clicks/number of turns.

(a) The average time between mouse clicks for subjects during agent strategy: Adapt, EQHare, EQStag, Random, and Win–Stay–Lose–Shift (WSLS). The data were not normally distributed; therefore subject performance against different agents was compared using Wilcoxon rank-sum tests (Bonferroni corrected, $p < .005$ was considered significant). Each mouse click indicated a desired movement on the game board or action taken to catch a stag or hare target performed by the subject. (b) The average number of turns taken by subjects during each agent strategy. The number of turns was taken cumulatively for all games in a particular strategy for each subject. The boxplots have the same notation as in Figure 5.

To test how quickly subjects were making decisions, the average time between mouse clicks and the number of turns taken by the subjects were analyzed (Figure 1.8; Table 1.5; Table 1.6). Subjects had significantly shorter delays in the EQHare condition for mouse clicks than all other conditions, and subjects had significantly longer delays in the EQStag condition when compared to the Random condition (Figure 1.8(a); Table 1.5).

Table 1.5. p-values for Wilcoxon rank-sum pairwise comparisons of average subject path deviation ratio in each condition over game losses.

	Adapt	EQHare	EQStag	Random	WSLS
Adapt		$< .0001$	$< .0001$	$< .0001$	$< .0001$
EQHare			$< .0001$	.6937	.0001
EQStag				$< .0001$	$< .0001$
Random					$< .0001$

Table 1.6. p-values for Wilcoxon rank-sum pairwise comparisons of average subject mouse click delays in each condition.

	Adapt	EQHare	EQStag	Random	WSLS
Adapt		$< .0001$	.3531	.0207	.2127
EQHare			$< .0001$	$< .0001$	$< .0001$
EQStag				.0009	.0279
Random					.2235

Subjects took nearly significantly longer between mouse clicks when playing with the Adapt agent compared to the Random agent, and EQStag had nearly significantly slower click times than WSLS. No other comparisons were significantly different. EQHare shows the most dramatic difference with a very short click time, indicating that in this condition, subjects had a target and trajectory clearly in mind for each game and made moves as quickly as possible. The increase in click time for Adapt might indicate that the subjects invested more time watching to see what moves the adaptive agent would make before the subjects made their movement decisions. Rank sum tests were also run for the differences between the average number of turns taken by subjects over all games in each of the five conditions (Figure 1.8(b); Table 1.6). EQHare was found to have significantly fewer turns taken when compared to all other conditions, followed by Random, which had significantly fewer turns taken than the remaining three conditions. EQStag was found to have nearly significantly more turns taken than WSLS. Subjects took more turns playing with the Adapt than with the Random agent. Both the increased number of turns and high mouse click delay indicate that the subjects were aware that the adaptive agent was not acting randomly and may show that the subjects attempted to guide the agent's behavior toward stag hunting to maximize payoffs.

1.3. Discussion

Economic game theory has had a long, productive history of predicting and describing human behavior in cooperative and competitive situations (Maynard Smith, 1982; M. A. Nowak, Page, & Sigmund, 2000; B. Skyrms, 2001). The theory of games has also been used to illuminate the neural basis of economic and social decision-making (Lee, 2008; Rilling & Sanfey, 2011). However, these studies typically have human subjects play against opponents with set-strategies and predictable behavior. By introducing agents with the ability to adapt to subject variation and the game environment, we were able to evoke stronger strategic variation in our subjects.

Specifically, subjects played the socioeconomic game known as the Stag Hunt because of its advantages for studying cooperation, teamwork, and social signaling (Skyrms & Pemantle, 2000; Skyrms, 2004). In a Stag Hunt, subjects must weigh the decision of hunting a valuable stag, which requires the cooperation of another player, against hunting a hare, a less valuable but more easily obtainable prey (i.e., cooperation is unnecessary). Because it has both a cooperative and non-cooperative equilibrium, as well as a temporal aspect (e.g., hunters can change their decision as the hunt progresses), the Stag Hunt may be a better model of cooperation and intention than the Prisoner's Dilemma, Hawk-Dove, or Ultimatum Game.

The adaptive agent was constructed based on a variant of the Actor-Critic model, which contained one Critic that learned the expected reward of an action and another Critic that learned the expected cost of an action. The model was similar to prior work in which a computational model of neuromodulation and action selection was developed based on the assumptions that dopamine levels were related to the expected reward of an action, and serotonin levels were related to the expected cost of an action (Asher, Zaldivar, & Krichmar, 2010; Zaldivar, Asher, & Krichmar, 2010). The dopaminergic and serotonergic systems have been shown to influence the evaluation of rewards and costs for future decisions respectively, and have a strong influence on social decision-making (Boureau & Dayan, 2010; Cools, Nakamura, & Daw, 2010; Krichmar, 2008).

The main findings of the present study involve the differences in subject behavior when playing with an adaptive model, as opposed to preset, predictable computer strategies and purely random strategies. We found significant differences in scores, deviation from a direct path to the desired target, delay between movement mouse clicks, and the ratio of stags to hares caught. It was found that subjects had more variation and uncertainty in their play with the Adapt agent. Additionally, close examination of the Adapt agent revealed that it not only altered play based on its own position on the game board, but also monitored the human players' relative locations on the board. Lastly, our findings indicate that there may be a divide in the subject pool that defines two distinct types of reactions to the adaptive model: those that become highly cooperative by primarily hunting stag with other players and those that become highly uncooperative by primarily hunting hare on their own.

Subjects playing with an adaptive agent may be investing more time and effort in trying to discover the agent's strategy, recognizing that a strategy was, in fact, being used rather than the agent taking random actions. As seen in Figure 1.8(b), subjects took significantly more turns when playing with the Adapt agent than the Random agent. This could indicate that either players were attempting to influence the agent's actions by executing guiding movements toward the desired target, or that players found it necessary to change their strategies mid-game, abandoning their first target to pursue a different target as the agent's actions became clearer. In further investigation of the guiding hypothesis, the data was analyzed to determine which player arrived at the stag first in the Adapt condition on average. This was decided by identifying the player who landed within one square of the stag first. Subjects arrived first in over 1/3 of the games, indicating that on many trials, the subject attempted to show the Adapt agent cooperative intention. Further supporting the idea of subject observation and strategizing was the finding that the adaptive agent was shown to cause somewhat longer delays between mouse clicks than the random agent (Figure 1.8(a)), indicating that subjects spent a longer time thinking about their moves with the Adapt agent than with the Random agent. This extra time was likely used either to estimate the pattern of the adaptive agent's moves in order to choose the best target, or to develop a strategy to guide the adaptive agent towards the desired target.

Subjects showed greater uncertainty and varied strategy in play with the adapting agent compared to other conditions. In Figure 1.5, the average scores for Adapt were significantly different from every other condition except Random. However, the wider variance of the quartiles in Adapt suggests that some subjects varied their responses, possibly in an attempt to shape the adaptive agent's actions. This conclusion is compatible with the interpretation of the results in Figure 1.8 because the extra turns and extra time spent considering possible outcomes in the Adapt condition may also be an attempt to influence the adaptive agent. The path deviation analysis further supports these claims. Subjects deviated from a straight path more when playing with the Adapt agent, as opposed to other agents (Figure 1.7), providing more evidence that the Adapt condition may encourage subjects to either change their strategy mid-game or that they attempted to use guiding moves to influence the adaptive agent's behavior. Again, the significant difference between Adapt and Random underscores the point that the subjects treated the adaptive agent as if the agent was using a complex strategy rather than acting randomly. Figure 1.7(b) shows an even more pronounced difference between Adapt and the other conditions when comparing only the games in which the subjects did not successfully catch their target and were beaten by the computer. This result is likely found because in any condition besides Adapt, when the subject loses a game, it happens quickly as the agent is simply heading straight for a hare target. The adaptive agent is not likely to simply rush to a hare target unless it has been trained to do so by a frequently uncooperative subject.

In the Adapt condition, the agent is able to change its mind in deciding what target it will pursue mid-game, meaning that the path to a target for the adaptive agent is not as clear-cut and may change. This indicates that more thought on the part of subjects was put into interpreting the movements of the adaptive agent than any of the other strategies. The analysis of path deviation conducted for the Adapt agent showed a slightly higher, but comparable average value to average human path deviation. The Adapt agent's path deviation behavior indicates that it is interpreting the players' positions on the board and using past payoff information to determine its best strategy on any given turn.

When considering the Actor state tables, it becomes clear that the adaptive agent was in fact able to learn when to hunt stag and when to hunt hare depending upon both the agent's position and the subject's position to either target (Figure 1.4). The closer the agent was to the hare or the further the agent was from the stag, the higher its probability to hunt hare. However, the adaptive agent also considered the state of the other player. The closer the human subject was to the hare and the further the subject was from the stag, the more likely the adaptive agent would hunt hare. There were many cases in which the Adapt agent did not demonstrate a clear strategy and switched its hunting goal mid-game. For example, when the agent or the subject was far away from the stag, the probability to hunt a particular prey was roughly at chance. This result could be improved upon in future experiments by allowing the adaptive agent to play more games with the subject, therefore providing the agent more time to learn and develop its state tables, or by training different agents off-line (i.e., playing non-naïve agents). For the sake of the length of this experiment, however, the number of games per condition was capped at 50, the threshold found in simulation at which the agent began to exhibit strong strategic biases.

Table 1.7. P-values for Wilcoxon rank sum pairwise comparisons of average subject turn counts for each condition.

	Adapt	EQHare	EQStag	Random	WSLS
Adapt		$< .0001$	.0327	$< .0001$	.3110
EQHare			$< .0001$	$< .0001$	$< .0001$
EQStag				$< .0001$	.0078
Random					.0030

The possibility of three distinct groups within the subject pool is suggested by the stag-to hare-catch ratio data of the Adapt condition (Figure 1.6(d)). About half of the subjects in the Adapt data seem to form clusters at the extremes of the distribution, indicating a bias toward exclusive stag-hunting or exclusive hare-hunting, while the remainder tended to switch between stag and hare catching (see peak in the middle of Figure 1.6d). In contrast, the ratio of stag-to-hare catching against Random agents was somewhat normally distributed with a peak in towards equal stag and hare hunting (see Figure 1.6b). This implies that playing with the Adapt agent evoked different responses in some subjects over others, either encouraging strong cooperation or strong competition. For comparison, Figure 1.6(a) shows the EQStag data and Figure 1.6(c) shows the WSLS data. Both EQStag and WSLS appear to be heavily biased towards exclusive stag hunting. In the case of EQStag, stag hunting is obviously encouraged by the fact that the agent hunts only stag. In the WSLS condition, if the subject beat the agent once at catching a hare target, the agent would attempt to hunt stag in the next game and would continue stag hunting as long as the subject is also hunting stags, which subjects playing to maximize their score should do as predicted by game theory. Accompanying these histograms, the equilibrium table (see Table 1.7) shows each individual subject's personal bias in hunting over those four conditions, implying that many subjects have tendencies to cooperate and compete in this context, and that some subjects were strongly influenced to change those biases when playing against an Adapt agent (e.g., see subjects 33, 36, 37, and 52 in Table 1.7).

The suggestion that two or more types of strategies can emerge among individuals when playing socioeconomic games is similar to conclusions found in Asher et al.'s Human Robot Interaction (HRI) study in the game of Hawk/Dove using an adaptive model (Asher, Zhang, et al., 2012). The conclusions drawn from their Acute Tryptophan Depletion (ATD) data indicated a division in their subject pool very similar to the divide found in the current experiment. Their subjects, when tryptophan-depleted, fell into one of two groups; either more cooperative or more competitive during games, much like the present study's subjects while playing against the adaptive agent. The present study is further comparable in that the Reward and Cost Critics used here resemble the serotonergic/dopaminergic systems inspiring the model in Asher's study.

Variation between individuals in socioeconomic games may be due to differences in dopamine and serotonin signaling (Bevilacqua & Goldman, 2011; Hyde, Bogdan, & Hariri, 2011; Loth, Carvalho, & Schumann, 2011). For instance, a variation of an upstream promoter region of the serotonin transporter gene (5-HTTLPR) has been shown to influence both behavioral measures of social anxiety and amygdala response to social threats in humans (Caspi, Hariri, Holmes, Uher, & Moffitt, 2010; Caspi, 2003; Hariri, 2002; Lesch et al., 1996; Young et al., 2007). Subjects carrying the short allele variant of the 5-HTTLPR outperform subjects with the long allele in an array of cognitive tasks and show increased social conformity (Homberg & Lesch, 2011). Polymorphisms in dopaminergic genes, including variable number tandem repeat (VNTR) polymorphisms in DRD4 and DAT1, have been associated with poor 'action restraint' and 'action cancellation' (Congdon, Lesch, & Canli, 2008; Munafò, Yalcin, Willis-Owen, & Flint, 2008). The prevalence of such polymorphisms in the human population suggests that there is an evolutionary advantage for this variability, such as optimizing competition or cooperation in different situations and investigating this variation in games such as the Stag Hunt may be promising.

Several simulation studies are pertinent to the present results. The cooperation aspect of game theory was also explored in studies such as Valluri (2006), where a variant of the Prisoner's Dilemma was used in a simulation with adaptive agents. The Prisoner's

Dilemma was altered such that cooperation was able to evolve, albeit against classical game theory predictions, by being iterated and sequential. This means that agents played games repeatedly against the same opponents, with the second player knowing the first player's action before deciding on their own action rather than both players making their actions simultaneously. A Q-learning algorithm controlled agents with a similar SoftMax function as the one used in the current experiment. Because this version of Prisoner's Dilemma was able to evoke cooperation in its agents, it is comparable to the Stag Hunt. The link between the sequential iterated Prisoner's Dilemma and the Stag Hunt is the ability to see intentionality before making an action. In Valluri (2006), the ability of the agents to reach cooperation was attributed to the sequential nature of turns rather than the traditional simultaneous action selection. In the version of the Stag Hunt used in the present experiment, players could see the path of the agent and choose their actions based on that knowledge. In this way, the present methods agree with this prior simulation study. In a study by Calderon (2006) using the Ultimatum Game, a simulation model of phenotypic plasticity was used in order to determine the evolution of cooperation in a population. The results showed that when plasticity was increased, cooperation was also increased in terms of the threshold for acceptance and the offer amount. Agents learned at the end of each game; proposers increased the amount they offered by one if their offer was accepted, and decreased their offer by one if it was rejected in the last game. The same alterations were made by recipients for their acceptance threshold. The games played were strictly one-shot, as the agents did not retain knowledge of whom they had played or what their previous payoffs were. In the Ultimatum game, cooperation is contingent on reaching middle ground in which the proposer and the recipient both agree on the division of the resource. Calderon found that in his control group, which did not exhibit plasticity, the relative fitness was higher than in the group with plasticity (2006). Although this result appears to be a strike against adapting agents, Calderon states that the reason this occurs is in any case where two individuals share a behavior, the agent who had that behavior innately will outperform the adapting agent due to the adaptive agent's initial learning cost. This comparison is very similar to the comparison of the EQStag and Adapt agents in the present study, as higher scores were achieved when playing against the EQStag agent. While the EQStag agent began at cooperative equilibrium, there was inevitably a large cost accrued in the learning period needed for the Adapt agent to learn cooperation, and the subject to adapt to the Adapt agent.

The results of the present experiment have brought up some interesting observations for future study. The individual subject strategy differences while playing with the adaptive agent suggests that there may be phenotypical variation influencing this behavior. Additionally, the unique response overall to the adaptive agent in comparison to set-strategy agents invites further exploration of the adaptive agent's ability to evoke a social response akin to that of playing against another subject. In a future study, these two observations will be explored through their neural correlates to, in the case of the first observation, distinguish a difference in brain activity between the two equilibrium players, and in the case of the second observation, show the difference in response between adaptive agent opponents and other human opponents. This study will both qualify and quantify the adaptive agent's effect on subjects seen in the present experiment.

1.4. Conclusion

The main goal achieved by the present study was to show that adaptive agents were in fact able to create a significantly different response in human subjects than that of set-strategy agents. Adaptive agents are useful for interacting in a game environment due to their unique ability to evoke complex and interesting results in human subjects while learning strategies of their own from both experience and subjects' behavior. Having the experiment situated in a game allows for a level of control and customization that is valuable when conducting experiments of any degree of specificity. Because of the unavoidable degree of unpredictability encountered when using exclusively human subjects, the level of control afforded by the use of an adaptive agent is also desirable. The secondary goal achieved by the present study was to create computer agents that were able to learn in real-time without deliberate feedback outside of the game environment and have those agents mimic human behavior enough for subjects to learn to trust and cooperate with them in a relatively short time span. The ability of the adaptive agent to evoke a more complex reaction in human players warrants study into the social effects of human-robot interaction using robots that are able to better emulate complex strategies humans would use in a game environment. Future research in the field of adaptive agents may lead to robot or computer interfaces that are more natural or sociable, providing a smoother transition of complex technology into everyday life. In addition, adaptive agents have the potential to add a heightened degree of realism to HRI, specifically for socially affective robots (Thomaz & Breazeal, 2008).

CHAPTER 2: Investigation of Biases and Compensatory Strategies in the Probabilistic Wisconsin Card Sorting Test

Social interaction and performance in game environments is comprised both of theory of mind and decision-making abilities. Knowing when to make the right decisions at the right times is a product of experience, instinct, and the ability to cope with uncertainty in a dynamic environment. Therefore, to accompany analysis of theory of mind, the investigation of decision-making ability in probabilistic environments is relevant to the study of subject performance in game environments because of its usefulness in deciding, with very few social cues, what another player is likely to do given a particular scenario. As with most cognitive processes humans conduct on a daily basis, decision-making is subject to bias, whether that bias comes from evolutionary shortcomings that were adaptive in human development, or from an interruption in the stream of sensory data the subject receives from its environment. In order to study these biases and how they affect human decision-making during games, the Wisconsin Card Sorting Test provides a well-established research tool to use as foundation. A description of this work can be found in Frontiers in Psychology (Craig, Phillips, Zaldivar, Bhattacharyya, & Krichmar, 2016).

2.1. Wisconsin Card Sorting Test

Dating back to 1948, David Grant and Esta Berg's Wisconsin Card Sorting Test (WCST) is a task that is commonly used in assessing the ability to "set-shift," or change one's way of thinking in the face of new goals or stimuli (Bishara et al., 2010; Grant & Berg, 1948). This task is useful in studying, modeling, and diagnosing disorders in higher-level processing areas of the brain such as the prefrontal cortex (Dehaene & Changeux, 1991; Lie, Specht, Marshall, & Fink, 2006; Milner, 1963; Nyhus & Barceló, 2009; Robinson, Heaton, Lehman, & Stilson, 1980; Rougier & O'Reilly, 2002). In the WCST, a subject is presented with one reference card and three to four choice cards. Each card contains an image with a particular shape, color, and number of items, and is designed such that each choice card's feature expressions are mutually exclusive. Every choice card matches a different feature of the reference card. In each trial, one feature is selected as the "rule," and the objective is to select a card that matches the rule for the reference card. For example, if the rule is green, the correct choice would be the card that contained green items, irrespective of the number or shape of items on that card. The WCST consists of several iterations of trials that use the same rule, followed by a rule shift that requires subjects to change their behavior.

With the goal of evoking and quantifying changes in strategizing, behavior, and biases due to uncertainty, we developed a modified version of the WCST called the probabilistic WCST (pWCST). pWCST incorporates an element of uncertainty in the form of a probabilistic rule selection, and an option to forage for information by observing a trial. Each trial has a set of three probabilities corresponding to the likelihood that a particular feature will be the rule. For example, the rule could be dictated by a 90% chance of shape, 7% chance of color, and 3% chance of number of items. These probability distributions are referred to as the Top, Middle and Bottom rules, respectively, and these base percentages are referred to as the "ground truth probability distribution" in the paper. Because humans are oftentimes shown to make irrational decisions regarding probabilistic assessment in the face of uncertainty, incorporating a varying degree of uncertainty into the WCST adapts the task into a tool that can be used to evoke and quantify the degree of change in behavior that is introduced into the decision-making process for unpredictable events. We hypothesize that by increasing the level of uncertainty in the WCST task, we will evoke biases and strategic changes in subjects that correlate to the degree of uncertainty.

Card-based tasks in the past have been common for assessing the patterns of balancing exploration and exploitation (Hoehn, Southey, Holte, & Bulitko, 2005; Sang, Todd, & Goldstone, 2011; Worthy, Maddox, & Markman, 2007), both for testing diminishing resources as would be experienced in real world explore/exploit tasks, and for understanding the underlying probability in action choices. The element of information foraging is commonly studied using decision-making experiments, often utilizing a probabilistic task, although to our knowledge, the WCST has not been previously modified to accommodate this mechanism. Introducing uncertainty into the WCST has been explored in a previous study by Wilson and Niv (2012). Wilson and Niv used the WCST in conjunction with a Bayesian model in order to examine the methods by which humans decide what information to learn in a changing environment.

The present study moves beyond Wilson & Niv's paradigm with the addition of "Observe" trials, an option to collect information without affecting one's score. While Wilson and Niv kept task uncertainty fixed, the present study investigates the effect of a variable level of uncertainty, which is hypothesized to affect both Observe behavior and strategy usage. The Observe feature introduced in the pWCST allows the subject to explore potential payoffs rather than exploit immediate gains. This adds an alternate option, similar to no-choice utility (Howes, Duggan, Kalidindi, Tseng, & Lewis, 2015), to the classical explore/exploit tradeoff where subjects can practice alternate strategies in a subsequent choice phase. Another similar task of note is the probabilistic lights task utilized by Navarro and Newell testing the theory that humans tend to assume a higher underlying rate of change than the ground truth probability distribution during a probabilistic task (2014). In this task, subjects were told to predict which of two lights would come on, given that only one would light up on each trial. In the dynamic condition, subjects were told that the bias on these two lights could randomly change, and in the static condition, the bias was to stay the same. In both conditions, the bias was always 70% and 30%, with a 1.6% chance of switching in the dynamic condition. Subjects were allowed to either Observe or Bet on each trial, with Observe allowing them to test their response without gaining or losing points, while Betting would result in a change in points. Results showed that in both static and dynamic conditions, subjects significantly overestimated the amount of switching that occurred, which confirmed the hypothesis and suggested that it was more costly to underestimate a rate of change than overestimate. Similarly to the pWCST, subjects were allowed to Observe trials without any point gain or loss, as an alternative to betting their real points. While Navarro and Newell's task adequately probed subject Observe behavior, the pWCST takes this task paradigm further with the inclusion of differing probability distributions, adding a level of uncertainty that cannot be directly tested in the probabilistic lights task. Additionally, the introduction of varying probability levels allows for the investigation of effects due to probability magnitudes, potentially identifying overweighting and underweighting effects, as well as differences in the level of Observe reliance.

2.2. Cognitive Biases

Cognitive biases are deviations from normative strategies, which occur both consciously and unconsciously in human decision making, to quickly and efficiently cope with uncertainty or task difficulty. While these biases can lead to non-optimal action decisions (Tversky & Kahneman, 1974) it has also been shown that such biases, under the right circumstances, can result in near-optimal task performance, designating them, not as irrational, but as "bounded rational" behavior (Gigerenzer & Goldstein, 1996). Therefore, these biases are interesting to study both for their commentary on shortcomings of human decision-making, as well as their insight into conscious and subconscious techniques that allow for fast and frugal yet high-yielding processes for creating action decisions. The decision-making behaviors of interest, including cognitive biases, heuristics, and non-optimal/bounded rational strategies, in the present paper include negativity bias, probability matching, and satisfaction of search, all of which the pWCST is expected to evoke, that would not be expected in the traditional WCST.

Negativity bias is the unbalanced increase in salience of negative over positive feedback (Carretié, Mercado, Tapia, & Hinojosa, 2001; Ito, Larsen, Smith, & Cacioppo, 1998; Rozin & Royzman, 2001; Vaish, Grossmann, & Woodward, 2008). Negativity bias may have adapted by allowing humans to focus more on information that was potentially harmful rather than helpful, as neglecting harmful information is more likely to shorten one's lifespan. However, the presence of negativity bias, through prioritizing avoiding negative behavior patterns, could potentially delay or prevent time spent on the development of positive behavior patterns. The present pWCST is hypothesized to evoke negativity bias while WCST would not, as pWCST is inherently more challenging due to the probabilistic rule, and in higher uncertainty levels, is expected to lead to higher amounts of negative feedback. This may result in subjects spending more time foraging for information when they feel they are receiving too much negative feedback resulting from their choices.

In probability matching behavior, an individual will perform actions that roughly mirror the underlying probability structure inherent in a task environment (Shanks, Tunney, & McCarthy, 2002; Vulkan, 2000; Wozny, Beierholm, & Shams, 2010). Probability matching can occur in situations where it is advantageous to explore options rather than exploit the best choice. When a subject feels uncomfortable with their ability to identify and exploit the most valuable option, subjects may revert to probabilistic search for information about the rules of their present task (Gaissmaier & Schooler, 2008). Probability matching is another suboptimal decision framework that pWCST is expected to evoke where WCST would not, as it is a behavioral strategy that is employed during situations in which the subject receives probabilistic payoff, a novel inclusion in the pWCST paradigm. We hypothesize that subjects will utilize probability matching as a means to cope with the uncertain environment in the pWCST, rather than the optimal strategy of continuously selecting the highest probability feature. During periods of moderate uncertainty, we expect subjects to use the Observe option to derive the "expected uncertainty" from a block of trials (Yu & Dayan, 2005).

In satisfaction of search, the individual possesses a threshold at which they determine they have collected enough information for their task (Fleck, Samei, & Mitroff, 2010; Simons, 2010). As confirmatory evidence is acquired, less evidence is required from the information foraging process, which is a heuristic that can save time and energy but is not inherently rational (Gigerenzer & Goldstein, 1996). Having preconceived notions about the underlying nature of a task is in itself a bias, and when those notions are accompanied by confirming evidence that is determined by a probability rather than a static metric, these biases become compounded. We hypothesize that, during higher uncertainty levels, the satisfaction of search threshold will be higher than in lower uncertainty levels. This may be due to a resulting increase of conflicting information necessitating a larger sample size to reach the same level of confidence. The pWCST allows for the opportunity to study satisfaction of search where the WCST would not, owing to the inclusion of an option that allows for risk-free information foraging. Through the incorporation of the Observe option, we predict that the satisfaction of search threshold (i.e., the number of Observe trials necessary for a subject to be confident enough in their rule beliefs to cease foraging) will be higher under higher uncertainty conditions.

Win-Stay-Lose-Shift (WSLS), a strategy commonly used in game theory (Imhof, Fudenberg, & Nowak, 2007; M. Nowak & Sigmund, 1993) involving staying with an action following its successful use and shifting to another action following its unsuccessful use, was selected for strategic assessment alongside probability matching during the pWCST. These strategies were chosen for comparative analysis in order to provide reasonable baselines for subject performance that realistically encompassed biases, suboptimal strategizing, and human limitations. While probability matching and WSLS behaviors are formal strategies that lead to positive performance in the pWCST, the pWCST is akin to a multi-armed bandit task (Michael D. Lee, Zhang, Munro, & Steyvers, 2011) coupled with the WCST as a result of the introduced probabilistic component. In the pWCST, optimal task performance consists of always selecting the highest probability feature.

2.3. Methods

2.3.1. Human Participants.

Sixty subjects (ages 18-25) were recruited in two sessions of 30 subjects through an online database maintained by the Experimental Social Science Laboratory (ESSL) at the University of California, Irvine (UCI). This database is comprised of UCI's undergraduate and graduate student population that have agreed to participate in computer experiments based in social and economic decision-making conducted by members of the School of Social Sciences and affiliated organizations. Subjects were not selected using background characteristics (age, race, gender) other than their student status. Subjects participating in the second session were prescreened to ensure that they had not previously participated in the experiment. Experimental protocol was reviewed and approved by the UCI Institutional Review Board, and informed consent was obtained from all participants.

2.3.2. Experimental Design.

Subject data was collected using desktop PCs within the ESSL. Prior to the experiment, subjects were instructed on the basic structure of the task, the payment system consisting of a baseline amount plus an incentive sum reliant on their performance, and their right to cease participation without penalty for any reason at any time. Subjects then participated in two behavioral tasks, the pWCST and a version of the Wason Selection Task, a similar decision-making task also investigating biases and compensatory strategizing related to uncertainty, in a randomly assigned order. The Wason task will not be discussed further in this paper, as it is the subject of a separate analysis. Each behavioral task incorporated a brief tutorial to train subjects on the tasks, as well as inform them that the rule involved in the tasks would change throughout the experiment. Subjects were not, however, informed that the rule sets they would encounter would be probabilistic, in an effort to preserve unprejudiced strategizing in the face of unreliable feedback.

Upon completion of both tasks, subjects received a $7 flat rate for participation, as well as compensation dependent on their performance in both tasks. The rate of compensation for WCST was $0.02 per point, with a minimum performance-based payout of $2. Total payments ranged from $10 to $30.

Table 2.1. Probability sets for WCST.

Uncertainty	Top rule	Middle rule	Bottom rule
No	100%	0%	0%
Low	90%	7%	3%
Moderate	75%	20%	5%
High	60%	30%	10%

Table 2.2. Block order/Criterion for WCST.

Uncertainty	Block order	Criterion
No	1, 7, 10	90% correct
Low	2, 6, 11	80% correct
Moderate	3, 5, 8	65% correct
High	4, 9	50% correct

2.3.3. Probabilistic Wisconsin Card Sorting Test (pWCST).

Subjects played a total of 550 pWCST games, which were split into 11 blocks of 50 games each. There were four block types – No, Low, Moderate, and High uncertainty – all of which were presented three times each except High, which was only presented twice as a result of the block order. The probability sets associated with each block type are listed in Table 2.1, and the order of the blocks is listed in Table 2.2. The first third of the condition presentation order was designed to scale up the uncertainty-based difficulty gradually in order to investigate the effect of increasing uncertainty on strategy choice, information foraging, and the appearance of biases. As stated in the hypotheses, it was expected that altering the underlying feature probabilities and increasing uncertainty would lead to an increase in information foraging and satisfaction of search threshold, as well as the emergence of negativity bias. The second third of the order was intended to assess the rate at which subjects would adapt to a gradually decreasing uncertainty under their previous expectation derived from the trend of increasing uncertainty, which was expected to reverse the hypothesized increase in information foraging behavior and other changes in formal strategizing. The final third of the condition order was used to assist in testing for order effects and to incorporate more unpredictable jumps in uncertainty level between blocks.

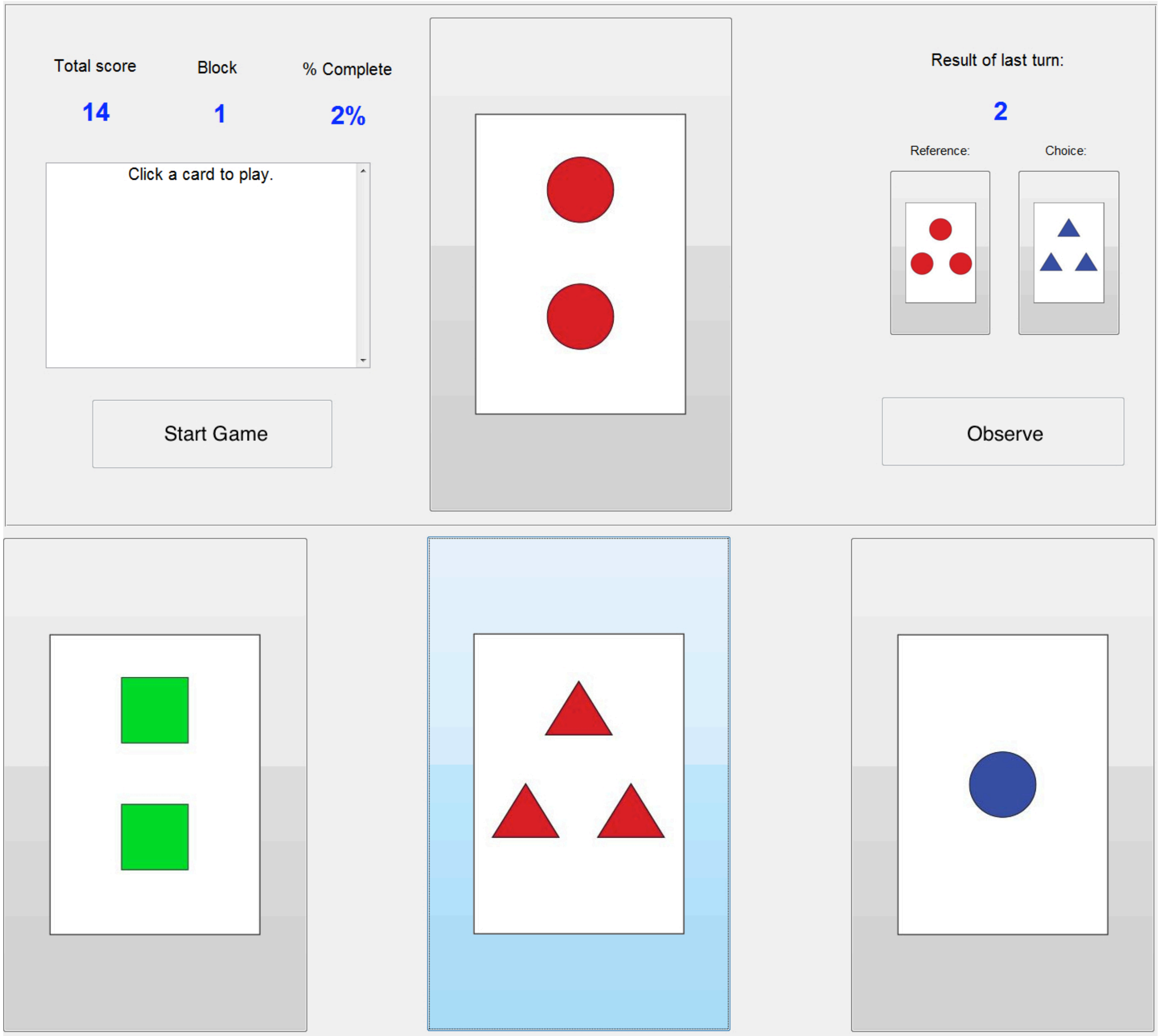


Figure 2.1. Screenshot of pWCST GUI.
On a given trial, subjects would choose between one of the three cards on the bottom, attempting to match the rule of common feature between the reference card and the target. Two points were received for each choice matching the rule, and two points were taken for any choice that did not match the rule on that given trial. Alternatively, subjects could click the Observe button in order to simulate a trial, in which a target card was randomly chosen and presented to the subject along with the score they would have received had they chosen it, with no actual point gain or loss.

Subjects were required to choose a feature (color, numerosity, or shape) of the presented cards that they believed to be in compliance with a particular rule (Figure 2.1). Each feature in a rule set had a probability associated with it. A feature set was randomly assigned to a probability at the beginning of the experiment and again each time a criterion for successful trials was met. This criterion was a percentage of correct answers out of the most recent ten trials (see Table 2.2). For example, in a Low uncertainty block, subjects who chose correctly on at least eight of ten trials received a rule change. To reduce the predictability of set shifts, assessments did not begin until a randomly selected number of trials (i.e., between ten and fifteen) had elapsed since the last block change or feature set shift.

In an effort to assess subjects' preference for foraging behavior to cope with the uncertainty of their task, subjects had the option to "Observe" rather than choosing one of the three cards during each trial. In this case, the subject did not select a card for that trial. Instead, by observing, a choice card was selected at random and the subject was informed what the card was and whether or not it followed the rule. By using Observe, subjects had a chance of collecting information about the rule until they were able to reduce their own level of uncertainty enough to select cards on their own. In essence, an Observe trial could be used to obtain information, but resulted in a loss of potential points. Alternatively, an Observe trial could be used to spare the subject from losses. The first block did not feature the Observe option, as it was meant to provide subjects with practice for the fundamental task and to determine whether they had fully understood the instructions.

The pWCST software interface (Figure 2.1) consisted of a reference card and three clickable cards that subjects chose from during each game. If a subject selected a card that matched the correct feature of the reference card, they received two points. Two points were taken away for each incorrect answer. The graphical user interface allowed subjects to see the percent of the task that they had completed, the block they were in, their total score, the reference and chosen card from the previous game, and the points received on the previous game. Subjects advanced through trials at their own pace, with the majority finishing both behavioral tasks in 60 to 90 minutes.

Unless otherwise specified, all reported p-values were derived using the two-sample Kolmogorov-Smirnov hypothesis test (refer to MATLAB kstest2). Because these p-values were based on multiple comparisons, the significance threshold was Bonferroni corrected by dividing 0.05 by the number of comparisons.

2.4. Results

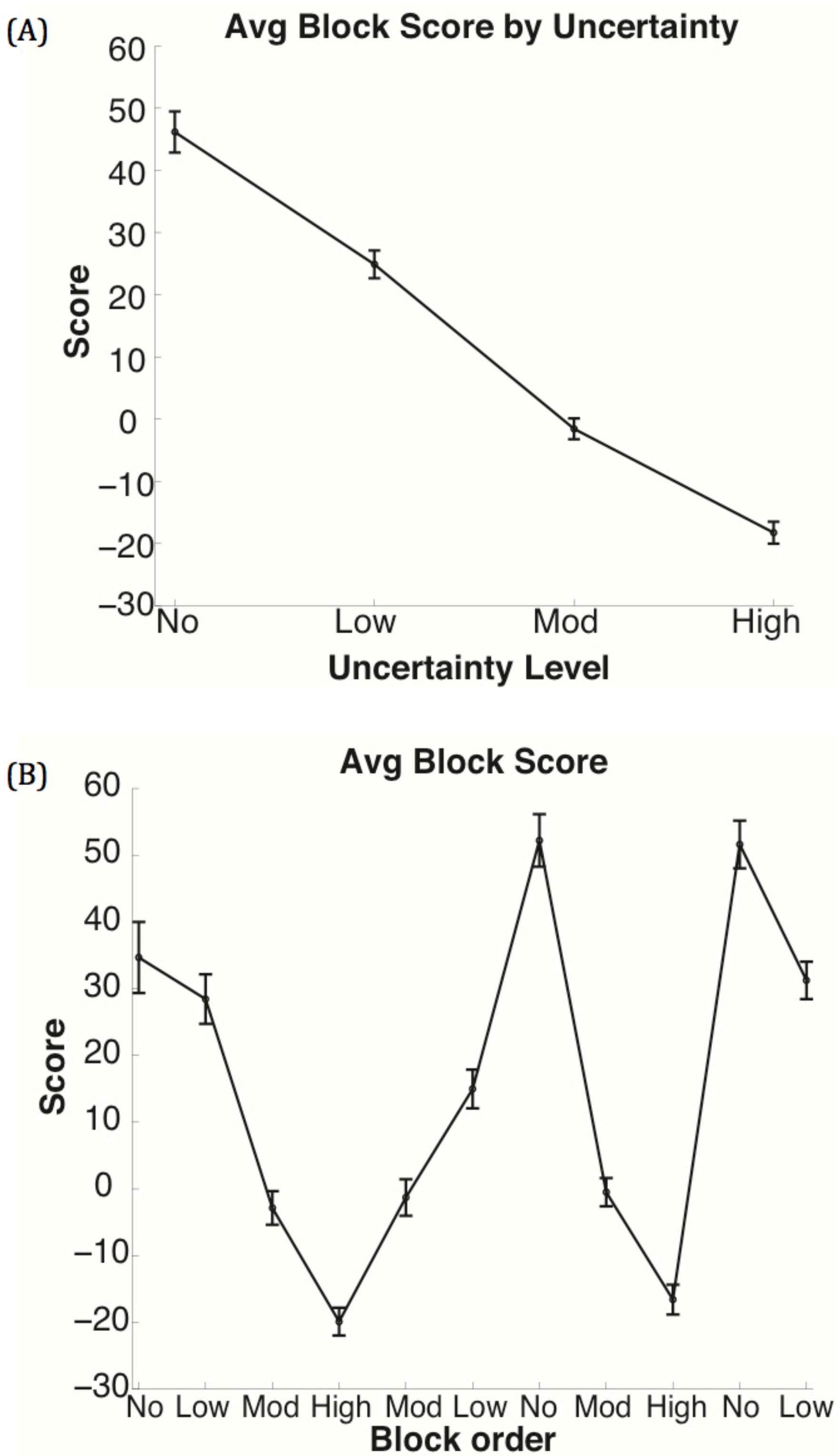


Figure 2.2. Average score per block by uncertainty.
Average score decreased as the block uncertainty level increased. In **(A)**, the x-axis is grouped by uncertainty level. In **(B)**, blocks along the x-axis are arranged in the order in which they were presented to subjects. Bars denote the standard error.

2.4.1. Score.

Subject performance varied depending on the level of uncertainty, as revealed by the average score per uncertainty level (see Figure 2.2**(A)** and Supplementary Tables 1-3) and average score per block (see Figure 2.2**(B)** and Supplementary Tables 4-6). The score per block averaged over all 60 subjects was significantly different between the four uncertainty levels ($p < .001$), with higher uncertainty associated with lower, often negative, scores.

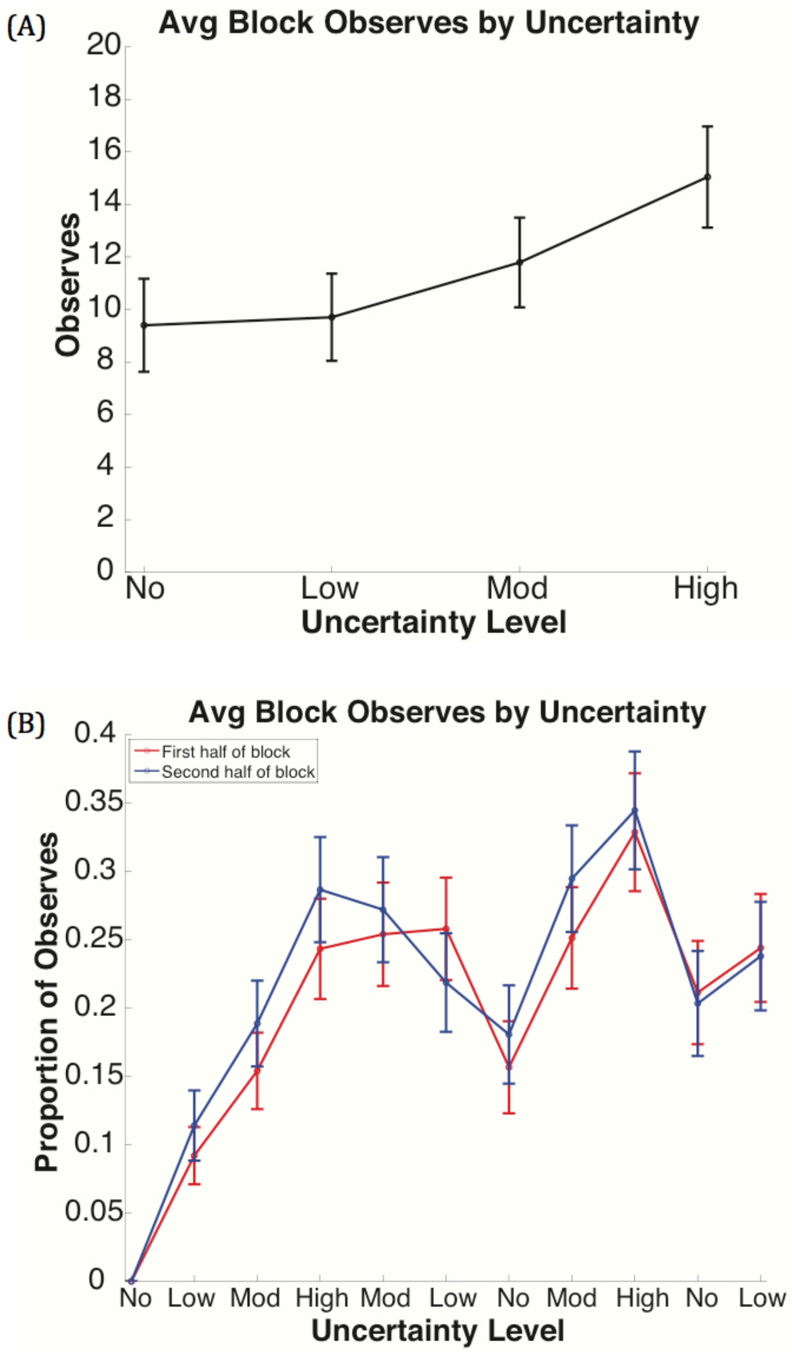


Figure 2.3. Average Observe use per block by uncertainty.
(A) Average Observe usage increased alongside increasing block uncertainty. **(B)** Observe usage per half-block by uncertainty. Observe usage largely increased during the second half of each block. This gap increased along with increasing uncertainty and decreased with decreasing uncertainty, in some blocks even causing usage in the first half to overtake the second half. In **(A)**, the x-axis is grouped by uncertainty level. In **(B)**, blocks along the x-axis are arranged in the order in which they were presented to subjects. Bars denote the standard error.

2.4.2. Observe Usage.

We analyzed Observe use in order to quantify how changes in uncertainty level could affect satisfaction of search thresholds. Search behavior, or information foraging, was assessed as use of the Observe option, a means of collecting information regarding the feature rule without risking points. Typical subject behavior was altered by changes in uncertainty level, as revealed by both the number of times the Observe option was used in blocks of each uncertainty level (see Figure 2.3 and Supplementary Tables 7-9). The number of trials in which Observe was used in a block increased with uncertainty level (see Figure 2.3**(A)**), and comparisons of No vs. Moderate ($p = .006$), No vs. High ($p < .001$) and Low vs. High ($p = .002$) uncertainty were found to be significant. To examine how subjects used Observe to gain knowledge about the task structure, we compared the amount of Observe trials during the first and second halves (25 trials each) of each block (see Figure 2.3**(B)** and Supplementary Table 10). Although no comparisons between the first and second halves of a block were shown to be significant, there was a trend indicating that Observe trials were more common in the second half than the first half of a block in most blocks, suggesting a subject preference to perform their own sampling within a new block before resorting to Observe use. This difference in Observe usage increased alongside increasing uncertainty and decreased with decreasing uncertainty. Taking the number of Observes per block as insight into information foraging behavior, these results support the hypothesized increase in the satisfaction of search threshold at higher uncertainty levels.

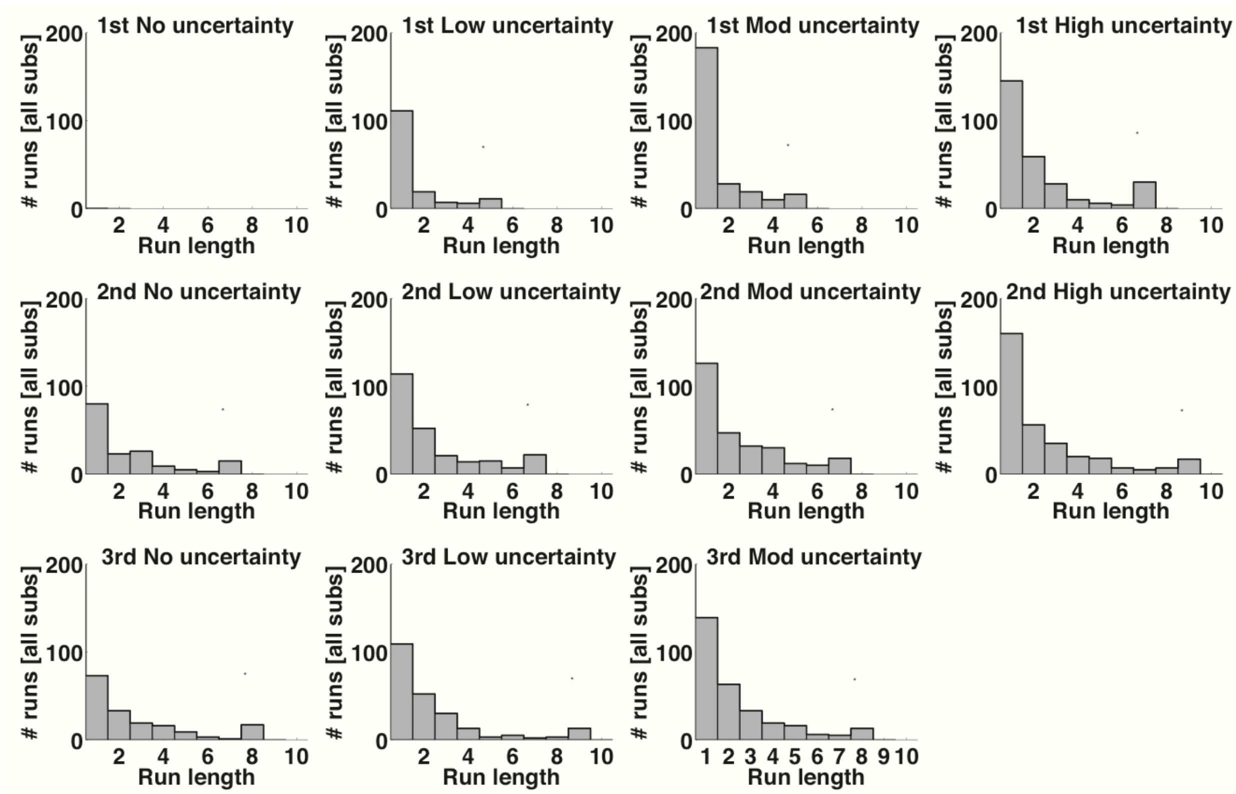


Figure 2.4. Run length of Observe usage by all subjects per block.
Blocks are grouped horizontally by presentation number of each uncertainty level, and vertically by uncertainty level. The title numbers indicate the true block order as experienced by the subjects. The length of the tail, and thus the average length of runs of Observe trials, increased both with increasing uncertainty and successive presentations of each uncertainty level.

To further measure how Observe usage changed over time and uncertainty level, we measured the runs of consecutive Observe usage (see Figure 2.4 and Supplementary Tables 11-16), which here is defined by the number of trials in a row a subject chose to Observe rather than picking one of the three cards. As shown in Figure 2.4, average number of Observe runs increased with both uncertainty (compare rows of Figure 2.4) and time (compare columns of Figure 2.4), although the only comparison that reached significance under a 2-sample Komolgorov Smirnov test was that of the first and second presentations of the Low condition (Blocks 2 and 6 in Figure 2.4) ($p < .015$). In accordance with the previously discussed results, this increase in runs of Observe usage also comments on satisfaction of search, which rose over time alongside score (see Figure 2.2**(B)**), suggesting more accurate performance during non-Observe trials based on a more extensive collection of information over the length of a block.

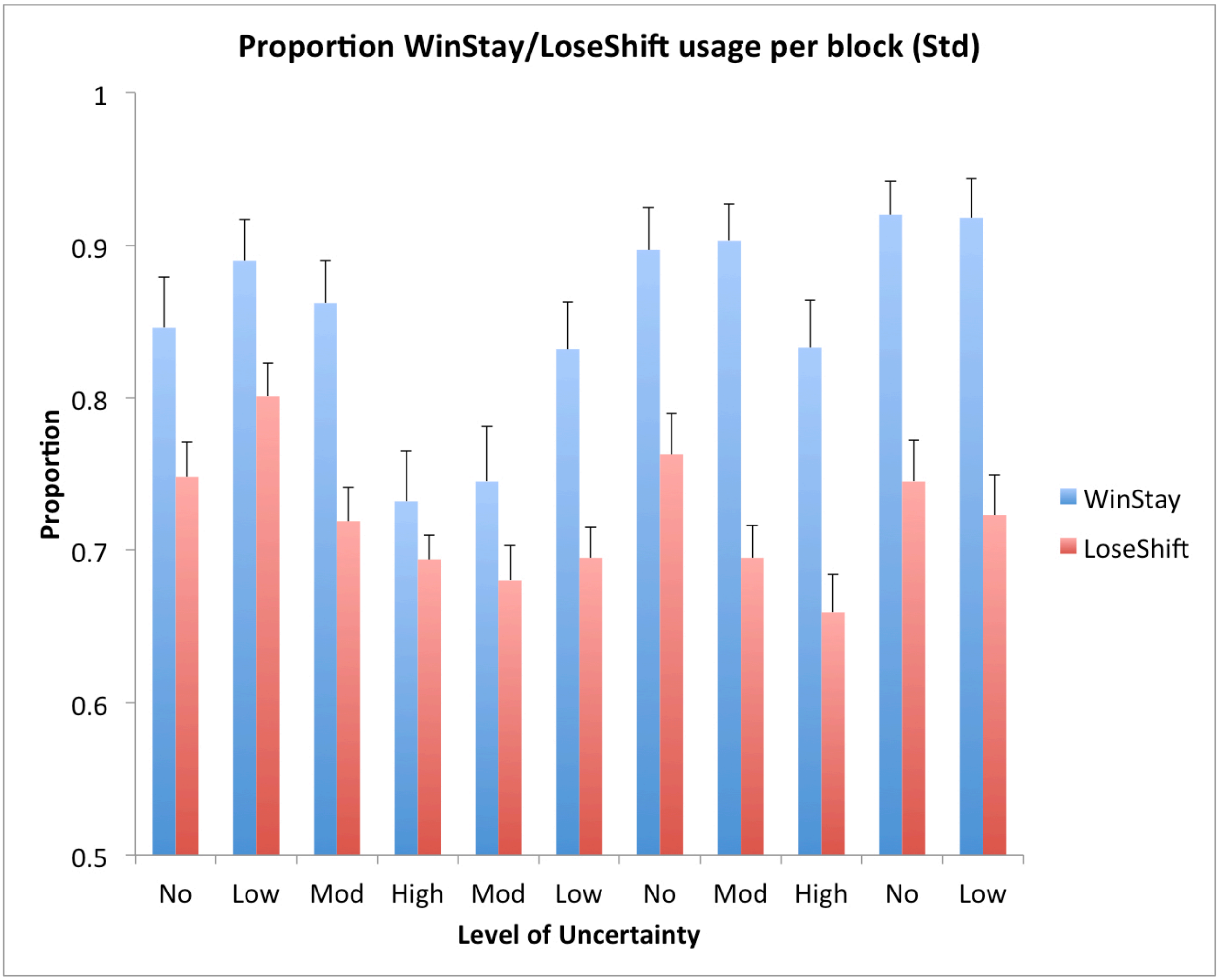


Figure 2.5. Win-Stay and Lose-Shift use by uncertainty.
"Win-Stay" bars denote the percentage of Win-Stay behavior out of all trials in which the subject chose the correct rule, while "Lose-Shift" bars denote the percentage of Lose-Shift behavior out of all trials in which the subject chose incorrectly. Bars denote the standard error and x-axis is arranged in the order the blocks were presented (Table 2.2).

2.4.3. Win-Stay-Lose-Shift (WSLS) Strategy.

WSLS usage was sensitive to the level of uncertainty and the order of blocks. The uncertainty level and the block order affected the use of Win-Stay and Lose-Shift behavior (Figure 2.5 and Supplementary Tables 17-28). The significant dominant strategy was found to be Win-Stay (i.e., choosing the same feature after a win), for win trials and Lose-Shift (i.e., choosing a different feature after a loss) for lose trials over all blocks ($p < .001$). While no significant order effects exist, a trend of increasing Win-Stay usage over time for all uncertainty levels was observed, excepting for the second Moderate and Low blocks. On average, Win-Stay and Lose-Shift strategy use decreased with increasing uncertainty, and this effect was significant for comparisons between No and Moderate, No and High, and Low and High uncertainty levels for both Win-Stay and Lose-Shift, as well as Low and Moderate Win-Stay (Lose-Shift Low vs. High: $p < .006$; all others: $p < .001$).

For some comparisons, blocks of high uncertainty were found to bias future behavior. The proportion of trials in which a Win-Stay strategy was used in the second Moderate and Low uncertainty blocks was lower than that of the other presentations of Moderate and Low uncertainty blocks (see Figure 2.5). This effect was found to be significant in the comparison between the second and third presentations of the Moderate uncertainty level ($p < .001$), and was expected as a likely consequence of the experienced high uncertainty in the High block persisting to devalue reliability of confirmatory evidence in proceeding blocks, coupled with the subjects' lack of knowledge that the uncertainty level would decrease rather that increase over time.

A similar order effect was found in Lose-Shift strategy. However, in contrast to Win-Stay, the proportion of Lose-Shift trials did not show as dramatic of a change over time between presentations of the same uncertainty level (Figure 2.5). There was a significant decrease in Lose-Shift usage from the first to second presentations in Low uncertainty blocks ($p < 0.001$). There was also a substantial, but not statistically significant, decrease in Lose-Shift usage from the first to the second presentations of the Moderate uncertainty blocks. This behavior, similar to patterns of usage for Win-Stay, appears to be a consequence of lowered informational reliability from the High uncertainty block.

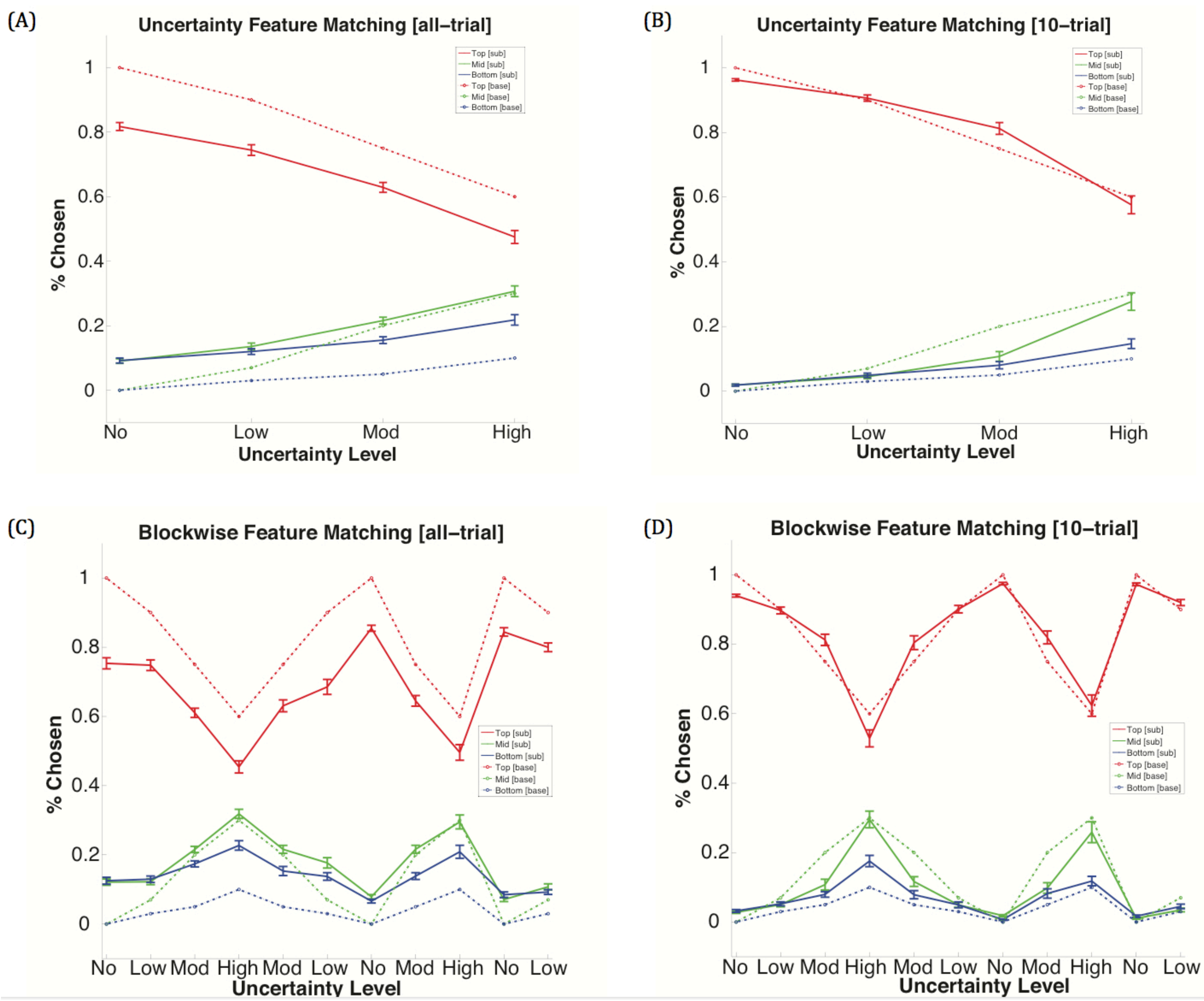


Figure 2.6. Subject feature matching vs. ground truth probability distribution averaged over all non-Observe trials for all blocks of each uncertainty level for all subjects.
Solid lines indicate subjects' choice percentages for the Top, Middle, and Bottom probability rules averaged over all trials of the No, Low, Mod, and High conditions. The dashed lines indicate the base rule frequency of the Top, Middle, and Bottom rules during each condition (see Table 2.1). (A) Feature matching percentages averaged over all non-Observe trials for each uncertainty level. (B) Feature matching percentages averaged over the last 10 trials (excluding Observes) before a rule shift (i.e. probability set to feature matchup scrambled) for each uncertainty level. (C) Feature matching percentages averaged over all non-Observe trials for each block. (D) Feature matching percentages averaged over the last 10 trials (excluding Observes) before a rule shift. Bars denote the standard error.

2.4.4. Probability Matching.

We analyzed subjects' choice behavior to see if they attempted to match the underlying distribution of rules rather than a normative strategy such as always choosing the most likely feature. Subject feature selection was averaged over each block to derive rule choice percentages. Feature selection percentages were calculated by dividing the number of times subjects selected the Top, Middle, and Bottom probability features by the number of non-Observe trials per block to obtain proportions (Figure 2.6(A) and (C), and Supplementary Tables 29-30). An additional analysis was conducted that used only the last 10 trials excluding Observes before a rule shift using the same metric, the window imposed by the rule shifting mechanism's threshold for analyzing correct responses (Figure 2.6(B) and (D), and Supplementary Tables 31-32). These data were compared with the ground truth probability distributions that were established prior to the experiment (see Table 2.1).

Results showed that the probability of selecting the Top, Middle, and Bottom feature roughly followed the ground truth probability distribution. Rather than choosing the highest percentage feature, Top rule selection decreased with increasing uncertainty.

To further examine the use of probability matching behavior, comparisons between uncertainty levels were made on the last 10 trials. Using a one-sample t-test, 10-trial uncertainty-wise comparisons showed that the Top rule was significantly different from the ground truth probability distribution only during No uncertainty ($p < .001$), the Middle rule was significantly different from the ground truth probability distribution during No, Moderate and the third presentation of Low uncertainty ($p < .001$), and Bottom was significantly different from the ground truth probability distribution only during the first presentations of No, Low, and High uncertainty, as well as the third presentation of No uncertainty ($p < .001$) (Figure 2.6(D)). In the 10-trial analysis, subject rule selection was only universally different from the ground truth probability distribution for all rules during the No uncertainty condition. Taken together, these results suggest that subjects did tend to probability match rather than use an optimal Top rule selection strategy.

2.4.5. Over/Under-selection.

Overselection, in the case of this experiment, is defined as choosing low probability rules at a significantly higher probability than the ground truth, and underselection is similarly defined as choosing high probability rules to a less often than the ground truth. We analyzed all trials and 10-trial block-wise data for similarity to ground truth. Using a one-sample t-test, all-trial block-wise rule selection proportions, with the exception of Middle/High, were significantly different from base rule percentages (Mid/Mod1: $p < .005$, Mid/Mod2: $p < .005$, Mid/Mod3 < 0.005; all others: $p < 0.001$) (Figure 2.6(C)). Subject usage of the Top rule universally fell under the ground truth, while selection of the Bottom rule always more frequent than the ground truth. Subject usage of the Middle rule fell over ground truth, but this difference declined with increasing uncertainty until the difference was not significant. These results provide evidence for underselection of the Top probability rule alongside overselection of the Bottom probability rule.

2.5. Discussion

In the present study, we showed that a probabilistic version of WCST with a means of information foraging is an effective tool for evaluating compensatory biases and suboptimal strategizing related to rational choice in economic decision-making. Through analysis of feature selection, it was clear that Win-Stay Lose-Shift and Probability Matching behaviors were the most prevalent strategies. The principal findings were that: 1) the threshold for *satisfaction of search* increased with uncertainty (Figure 2.3 and 2.4), 2) *negativity bias* occurred in trials following periods of high uncertainty (Figure 2.3 and 2.5), and 3) while subjects followed the trends of *probability matching* behavior, subjects also persisted in underselecting the Top probability rule while overselecting the Bottom probability rule (Figure 2.6). These findings address the relationship between uncertainty and the prevalence of biased and suboptimal selection behavior, a field with potential applications in the development of cognitive technologies and improving the decision-making techniques of human operators.

Through the incorporation of information foraging using an Observe option, satisfaction of search bias was demonstrated in the present study. The use of the Observe button increased with increasing uncertainty, indicating that the threshold for a subject's satisfaction of search (Fleck et al., 2010), or the amount of information a subject must collect before they are confident enough to take action in their current circumstances, similarly increased as hypothesized. This increase in Observe usage can be explained both as an avoidance tactic for losing points when the rule has a high degree of ambiguity, and as an attempt to learn a probability rule by collecting more information under conditions of high uncertainty. The relatively low Observe usage in lower uncertainty blocks may be the result of a lower satisfaction of search threshold or high confidence level. When analyzing the length of trials in which Observe was used consecutively, referred to as a "run," number of runs increased for both uncertainty level and block order (Figure 2.4). This trend also supports the increase of a subject's satisfaction of search threshold with higher uncertainty levels, but also indicates that subjects may be attempting to prevent further losses later on in the game by Observing more. By increasing the number of Observe runs in later blocks, subjects appear to behave according to an exploration paradigm in order to ensure they have the correct rule at multiple time points before risking their point reserve by participating in trials.

In the analysis of the prevalence of Observe usage during the first and second halves of a block, it was revealed that Observe usage was typically higher during the second half of a block, although this trend reverses in trials that follow presentations of High uncertainty blocks (Figure 2.3**(B)**). It is likely that the normal trend speaks to the subjects' preference to attempt to *exploit* their own theories in a new block before resorting to Observe usage to *explore* other options. The exception that occurred within the second presentations of Moderate and Low uncertainty blocks was likely a holdover following the first presentation of the High uncertainty block. Subjects did not yet know that uncertainty could decrease over time, and may have kept foraging for information in preparation for a higher uncertainty future. This behavior is indicative of an increase in risk aversion that led to an artificially increased level of satisfaction of search in lower uncertainty blocks as a result of the previously experienced higher uncertainty blocks. Furthermore, this result corroborates the observation that humans tend to shift to an information-seeking strategy when considering longer horizons for overall rewards (Wilson, Geana, White, Ludvig, & Cohen, 2014).

Negativity bias, or the tendency to remember negative feedback more strongly than positive (Rozin & Royzman, 2001), can be seen through increased risk-avoidance behavior that appeared after an increase in negative feedback. For example, the relatively slow (two block) return to a previously levels of WSLS following the first difficult, High uncertainty block suggests a negativity bias (see Figure 2.5) in that subjects' decisions were less likely to be influenced by confirmatory evidence and consistently select reinforced feature choices. Furthermore, this bias led to a substantial increase in Observe usage between the first and second presentations of Moderate uncertainty and a substantial decrease in score between the first and second presentations of Low uncertainty (see Figure 2.3**(B)**). As discussed in regard to satisfaction of search, the increase in Observe usage for higher uncertainty conditions supports the hypothesized evocation of negativity bias, as subjects may have opted for observing trials after experiencing a large degree of negative feedback as a result of their choices (see Figure 2.3**(A)**). Additionally, the finding that Observe usage is higher in the second half of a block other than in the two blocks directly following the first presentation of the High uncertainty condition is in support of a negativity bias (see Figure 2.3**(B)**). After experiencing a large degree of loss during the first half of a block, a subject may have made use of the Observe option more often in the second half of the block to prevent more loss. It might be that the feeling of loss persists from the end of the High uncertainty block through the beginning of the following blocks, creating a desire to prevent further losses until the underlying probabilities of the new block are better understood. Although this trend was observed, it would be of interest to conduct a follow-up study that utilizes more blocks in order to see the shift from high to low uncertainty enough times to confirm this theory with a higher significance level.

A trend indicative of over- and under-selection emerged regarding subject choice percentages for the Top, Middle, and Bottom probability features (Table 2.1) as a factor of uncertainty level. Subjects tended to overselect the Bottom feature, and underselect the Top feature across all uncertainty levels at a relatively consistent rate (see Figure

2.6). Subjects likely experienced trials in which the Bottom feature appeared commonly enough in the small sampling of trials to lead subjects to perceive their frequency as higher than the base truth. While the explanation for this result is not unequivocally clear, there are a few potential factors that may, by themselves or collectively, have resulted in the trends seen in the data. It is possible that an effect such as representativeness, the tendency to underselect and overselect as a result of small trial size and uncertainty (Tversky & Kahneman, 1974), could have led to this result. The most conservative explanation is that this effect was caused by noise in the data (Costello & Watts, 2014; Erev, Wallsten, & Budescu, 1994; Hilbert, 2012), perhaps due to feature choice error or the random sampling of features in lieu of a more concrete strategy. However, due to the consistency of this effect between subjects, it is unlikely that the identified effect could be completely described by noise. Moreover, if the effect was primarily due to noise, the deviation from the ground truth probability at high uncertainty levels should have led to increased random searching, which was not the case. Building slightly from this conservative explanation, it is possible that the uncertainty itself is the reason for the trend, given that a higher level of uncertainty leads to a longer feature sampling period as it becomes more difficult to identify the Top rule while experiencing a high degree of information obfuscating belief formation.

As an intermediary view between noise and bias, Gigerenzer posits the idea that, while the apparent trends in the data exist, they do not necessarily signify violations of probabilistic reasoning (Gigerenzer, 1991). In a similar view posed by Haselton, such trends are characterized as the result of “design features” rather than “design flaws” (Haselton, Nettle, & Andrews, 2005). Under this notion, the tactics exhibited by the subjects instead identify fundamental properties of probability and statistical theory in a situation involving varying degrees of uncertainty in task feedback. Building upon the ideas of Herbert Simon (1972), Gigerenzer provides additional applicable work concerning the idea of "bounded rational" strategizing, or the use of strategies that can be considered rational within the confines of a task. Given the uncertainty inherent in the pWCST, it is a further possibility that the identified trend falls under the category of a bounded rational strategy (Gigerenzer & Goldstein, 1996). As discussed below, one of our future plans is to develop an optimal model that could be compared to subject behavior in the pWCST. Such a model could provide more concrete support for one of these theories in explaining under- and over-selection of rules under uncertainty.

Perhaps the most interesting aspect of the subject choice percentage data is the trend of the Middle probability feature. Initially, the Middle probability feature, much like the Bottom probability feature, is overselected. However, as the uncertainty level increases, this overselection gradually tapers off until the selection percentage nearly matches the ground truth. It is possible that this result speaks to a threshold at which a probability becomes just substantial enough that it no longer falls prey to the bias that causes it. We predict that, in the pWCST paradigm, that threshold would fall at 33%, the uniform distribution given three feature choices. In the current paradigm, the Middle feature fell just short of that in the High uncertainty condition at 30%, and as would be expected, averaged subject selection of this feature was just above the ground truth probability distribution in this uncertainty level. In order to reveal more concrete evidence for this theory, in future work, we would like to add another uncertainty level to this paradigm in which all three features are at uniform probability. It would also be helpful in the future to design a model capable of playing the pWCST in an effort to make a firm assessment of baseline performance within each probability level, given the ambiguous nature of this particular result.

Lastly, probability matching behavior (Wozny et al., 2010) was observed in the analysis of overall strategy usage. Subjects feature choices roughly mirrored the ground truth probability distribution in trend in the 10-trial analysis (Figure 2.6 (B)), with a shift for over/underselection in the all-trials analysis (Figure 2.6 (A)). This finding indicates that rather than utilizing the optimal strategy of always selecting the Top probability feature, subjects favored selecting the features based on what they knew of their underlying ground truth probability distribution. An alternative explanation is that the variability seen in the data is caused by feature exploration, as subjects will need to test hypotheses by switching between the features before deciding upon their most rewarding rule. In order to address this possibility, we analyzed the probability matching data for just the last 10 trials before a rule shift occurred, with the rationale being that the subject needed to reach a high percentage of correct feature choices in order to engage the rule change, suggesting that they had at that point solidified their strategy. Even in the 10-trial analysis, not only did subjects still eschew the optimal strategy in favor of probability matching, but also their choices more closely matched the probability matching strategy than with the overall data. This finding provides strong evidence for the presence of this cognitive bias under varying levels of uncertainty.

There is the possibility that some of the biases found in this paper, especially the over/underselection of rules, may more conservatively fall under the explanation of noise in the data (Costello & Watts, 2014; Erev et al., 1994; Hilbert, 2012). While the explanations presented above serve as potential causes for the trends observed, we recognize the possible effects of noise or other potential explanations for the consistent trends in behavior across subjects leading to potentially suboptimal decision-making. While we cannot know the exact strategy, motivation, and perceptual acuity the subjects exhibited during these blocks, the trends of irrational and occasionally detrimental gameplay strategies that are evident in the collected data suggest the explanation that to some degree, subjects were under the influence of biases and suboptimal strategizing that prevented them from behaving rationally as dictated by the tenets of Game Theory (Zagare, 1984). However, the not yet unambiguated nature of these conclusions invites further investigation into the underlying causes for the patterns of behavior exhibited by subjects performing the pWCST.

The present study introduces a variation of the well-known WCST to examine biases, strategy usage, and decision-making under uncertain conditions. This study sought to expand upon previous experiments using the WCST and similar decision-making tasks to investigate uncertainty-related changes in behavior in subjects. The two primary extensions to the WCST are the introduction of uncertainty in the form of probabilistic feature selection, and the option to "Observe" a trial by allowing the computer to select a card for the user, showing them the outcome with no change to their score. These initial findings suggest that the present pWCST can evoke interesting deviations from normative behaviors.

The results of the pWCST task support the assertion that under increasing degrees of uncertainty, people tend to respond with a decreasing capacity for optimal decision-making behavior. However, there are a few modifications to the present paradigm going forward that would add to its statistical power and investigative scope. In future experiments using this task, it would be desirable to query further self-report data in order to elucidate the subject's mental state when performing the pWCST to form stronger conclusions regarding the reasons behind the biases that were evident in the data. The lack of self-report data is a limitation of the current study, and stronger conclusions might have been formed regarding biases had these data been collected. As mentioned above, we were limited by the number of uncertainty levels and recommend that future studies investigate additional uncertainty levels that staircase down to a uniform probability distribution, 33/33/33, in order to test the hypothesis that over/underestimation has a threshold, perhaps adding a shallower gradient between uncertainty levels for more accurate identification.

In regard to the discussion of over- and under-selection of features, the strategies taken by our subjects might be better understood if compared with an optimal computer model that played the pWCST. This is something we plan to explore in the future. Such a model would provide a better baseline for human subject performance in an effort to better assess the validity of claims about subject strategizing in the present study and might support conclusions about the presence of cognitive biases. Using performance data from the model, we would be able to comment on how each bias behaves in isolation, in the presence of other trends and biases, and the effects it would have on memory for future decision-making. Additionally, we would be able to distinguish, depending on whether the model outperformed subjects or roughly approximated their results, whether human performance on the pWCST could be considered bounded rational (Gigerenzer & Goldstein, 1996) or suboptimal. Another further analysis that could provide insight into the validity of cognitive biases in the data is minimum description length modeling (MDL) (Rissanen, 1983). Typically used in order to provide support for one theory among many that exist to encapsulate a trend found in data, MDL would be well-suited for disambiguating the aforementioned results, particularly regarding the over and underselection of features, which has been shown to have varying potential explanations.

An additional bias that might fit within the scope of the paradigm is confirmation bias. Confirmation bias is the practice of seeking out information that confirms one's prior beliefs rather than testing disproving information in an attempt to elucidate the ground truth in a situation (Doll, Hutchison, & Frank, 2011; Nickerson, 1998). While the confirmation bias may have been an adaptive shortcut that increased human survival by enabling expedient development of heuristics, this bias can also lead to either incomplete or incorrect perceptions of the world in many circumstances (Klayman, 1995). We have previously shown that satisfaction of search can work in conjunction with the confirmation bias to lower the threshold at which a subject stops foraging for information (Phillips, Chelian, Pirolli, & Bhattacharyya, 2014). This bias could be tested within pWCST by the addition of the ability to choose which feature is being used in Observe trials like in the paradigm utilized in Navarro and Newell (2014), instead of that feature being randomly selected. Confirmation bias would be a natural extension for this paradigm because, unlike the WCST, the pWCST, especially at high levels of uncertainty, would lead subjects to believe that low probability features are more common than they are as a result of conclusions based on a small sampling of trials.

Going forward, pWCST serves as a suitable platform for continuing to investigate biases and suboptimal strategies that are not commonly evoked by the traditional WCST, such as those described in this paper, as well as confirmation bias given the alterations described above. pWCST also holds potential use in investigating the tradeoff between balancing of information foraging and trial and error strategizing. The pWCST in its current form allowed for understanding when subjects were foraging for information or testing their hypotheses of the task structure. The revised Observe mechanism proposed above would make the held beliefs of the subject clearer, allowing a tighter investigation of when and how information foraging transpires. Gaining a clear understanding of the ways in which humans engage in suboptimal strategizing and the mechanisms that cause them to arise holds importance in a variety of applied positions, such as reducing human operator error, improving adaptive educational software, and modeling cognitive processes for medical and research applications.

CHAPTER 3: Theory of Mind-Inspired Adaptive Agents Utilizing an Actor-Critic Model with Forward Planning

In this chapter, we describe a study that investigates the relationship between autistic traits in typical individuals and implicit ToM during socioeconomic game play. Subjects played the repeated spatiotemporal Stag Hunt game with five different agents ranging from simple fixed strategies to adaptive strategies that simulated aspects of ToM. We quantified the level of autistic traits in our subclinical population using AQ scores from self-report survey responses (Simon Baron-Cohen et al., 2001). We hypothesized that subjects with high levels of autistic traits would engage in less cooperative strategies as compared to subjects with low levels of autistic traits. This disinclination to engage implicit ToM should be most apparent while playing with adaptive agents, conditions in which eliciting cooperative behavior from the simulated agent yields optimal rewards. In addition, we hypothesized that subjects with high levels of autistic traits would have difficulty understanding the intention of agents with ToM attributes, which may be exhibited through differences in visual displays of intention during game play. These displays of intention are indicators of cooperative or competitive behavior and therefore should follow the previously discussed correlation between cooperative preference and AQ. This work builds upon the materials and results discussed in Chapter 1.

3.1. Methods

3.1.1. Human Participants.

113 subjects (ages 18-33) were recruited in four separate sessions through the

Experimental Social Science Laboratory at the University of California, Irvine. This lab maintains contact information for a large population of undergraduate and graduate students that have agreed to participate in social and economic studies in exchange for monetary compensation. Subjects were not screened for characteristics (e.g. ethnicity, gender, age) beyond their current student status. The experimental protocol was reviewed and approved by the UCI Institutional Review Board, and informed consent was obtained from all subjects. Nine subjects were removed for failure to complete the AQ survey or the experimental task with full comprehension; the remaining 104 are included in the following analysis.

Figure 3.1. Stag Hunt game environment.
Subjects moved their avatar (stick figure) across the 5x5 game board using the arrow keys in order to catch a low payoff hare individually, or attempt to catch a high payoff stag with the computer agent (silhouette in computer screen). To catch a stag, both the subject and the agent must move their avatars into a square directly adjacent to the stag square, and the subject must

press the Space bar. To catch a hare, the subject must move their avatar on top of a hare square and press the Space bar. When a prey was successfully caught by one of the subjects, the game ended and points were awarded based on prey.

3.1.2. Stag Hunt Software.

In Craig et. al (Craig et al., 2013), software was developed to create a computer game version of the Stag Hunt. The game board (Figure 3.1) included a 5x5 grid of squares on which a human player and computer agent each controlled an avatar. The game board also included hare squares positioned at the leftmost and rightmost squares of the middle row of the board, and a stag image that was placed on a random square at the beginning of each game (but never appeared directly adjacent to a hare). The player avatars started on a random square on the game board such that they did not appear directly next to a hare, a stag, or each other. Two hares were included in order to prevent one player from having a significant location advantage derived from the random placement.

The human player moved their avatar using the left, right, up, and down keyboard keys, and caught a stag or hare by pressing the space bar. Diagonal movement was not allowed for either player. The human player was able to move as often as they liked, while the agent moved at randomized intervals between 300 and 900 ms, timed to reduce predictability and to approximate the range of reaction times exhibited by humans during pilot testing.

A successful hare capture required the player to be on the same square as the hare and to press the space bar, which awarded the first player who initiated a capture one point. A successful stag capture was worth 4 points to each player and required both players to be on squares adjacent to the stag square. The human player had to press the space bar on the computer keyboard when the agent was positioned adjacent to the stag and intended to hunt a stag. Successfully capturing the hare or stag ended the game. If neither the stag nor the hare was captured within ten seconds, the game timed out and no players received points.

Table 3.1. Agent types

Agent	**Influenced by Other Player**	**ToM**	**Planning**
ToMPlan	Yes	Yes	Yes
Plan	Yes	No	Yes
AC	Yes	Yes	No
Random	No	No	No
WSLS	Yes	No	No

3.1.3. Computer Agents.

Participants played the Stag Hunt with five simulated agents (see Table 3.1): Theory of Mind Planning (ToMPlan), Planning (Plan), Actor-Critic (AC), Random, and Win-Stay-Lose-Shift (WSLS). In Table 3.1, agents are organized by whether or not their behavior is influenced by the subject, whether or not the agent simulates ToM, and whether or not the agent uses planning to make action decisions. Subjects played multiple successive games (blocks) with the same agent type, allowing the agent to develop game outcome-based contingent strategies.

3.1.3.1. Fixed Strategy Agents.

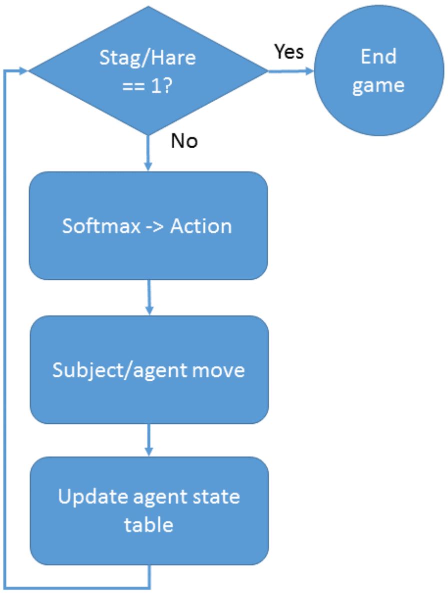
There were two fixed strategy agents: Random and WSLS. The Random agent was randomly assigned a prey (stag or hare) at the onset of each game and moved directly toward that goal without deviation. The WSLS agent used a fixed strategy in which its prey was selected based on the outcome of the previous game (with a random prey initially assigned on the first game of each block). If the agent won the previous game, it would select the same prey in the following game, and if it lost the previous game, it would switch to the other prey.

3.1.3.2. Adaptive Agents.

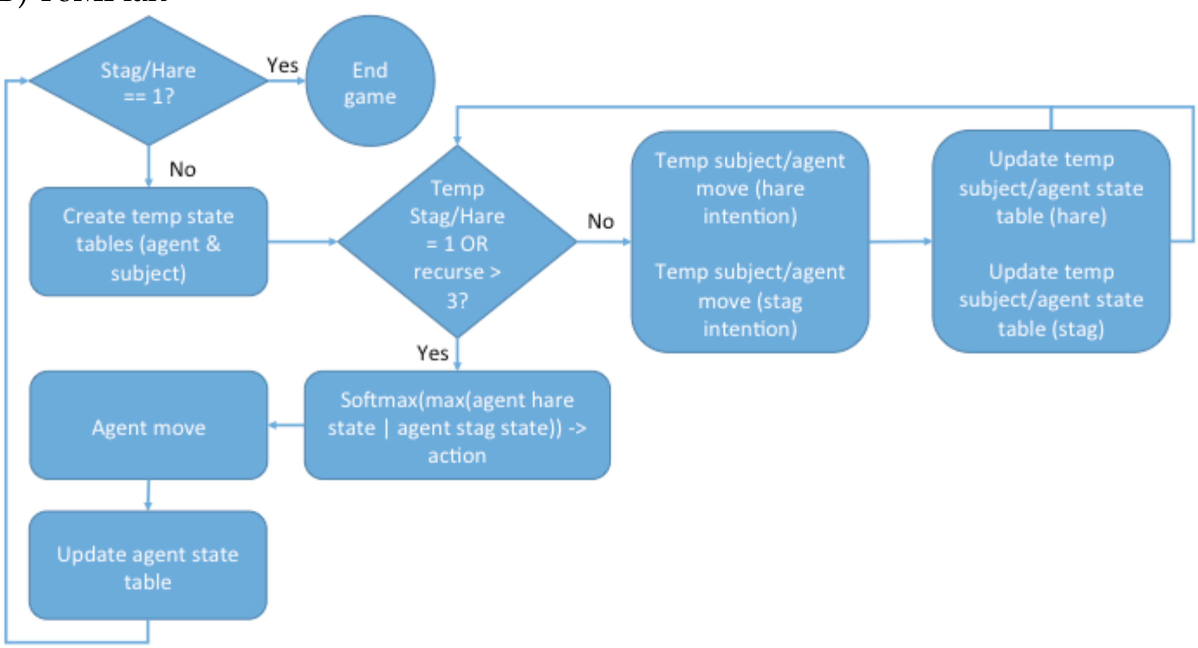
The remaining three agents used variations of an actor-critic model (Dayan, 2001), which was based on our previous work (see 1.2.3). The actor-critic model was comprised of state tables associating all possible configurations of the game board with their expected value under hare or stag hunting. These values were based on the agent's experience and were updated in real-time after the subject or the agent moved. Given the size of the game board (see Figure 3.1), the maximum distance of a player to a stag was four steps, and the maximum distance of a player to the nearest hare was three steps. When including the possibility of being on top of the target, this equated to 400 possible configurations, although not all of these states were reached during games. The actor-critic model included three state tables associated with these configurations: Reward, Cost, and Actor. The Reward and Cost tables held values of the expected reward and cost at each state, respectively. The Actor table combined this information together using a delta rule to provide a value that, through a softmax function, gave the probability of the agent to hunt each prey type at each state. This probability was used with a randomly generated number between 0 and 1 to make an action decision for the configuration of the game board when the agent made a move. If hare was selected, the agent moved one square closer to the hare, and if stag was selected, the agent moved one square closer to the stag.

Three agents used variants of the AC model: The actor-critic agent (AC), the theory of mind and planning agent (ToMPlan), and the planning agent (Plan).

A) Actor-critic



B) ToMPlan



C) Plan

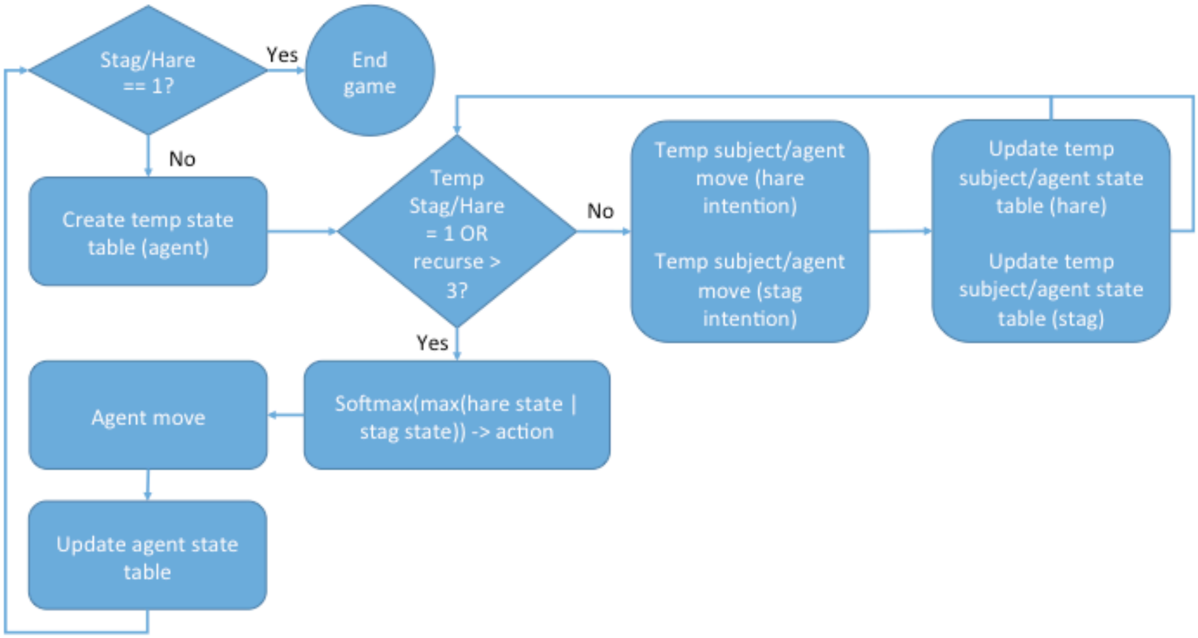


Figure 3.2. Flow charts of A) the actor-critic (AC) model, B) the theory of mind planning (ToMPlan) model, and C) the planning (Plan) model.
Model A) served as the foundation for all three models. On each turn, if an endgame was not reached through a stag or hare capture, the model would decide upon an action of either moving closer to the hare or the stag using state table information run through a softmax function. The agent would then move one square on the game board, update the state tables, and iterate again. In models B) and C), while an endgame state was not reached, a temporary copy of the state tables was made and a sample game for both hare and stag outcome was simulated until either endgame state was reached or three levels of recursion were completed. The prey with the simulated game containing the highest valued end state was selected, and this state was run through the softmax function to return a prey to act upon. While model B) held state tables for both the agent and the subject, model C) only held state tables for the agent, and those state tables did not keep track of any subject location information.

The AC agent used the basic actor-critic model described above (also see Figure 3.2A). This agent was able to learn in real time using the other player's behavior, but did not possess the ability to plan ahead.

The ToMPlan agent extended the basic actor-critic model to mimic mental simulation and planning (Figure 3.2B). It maintained real-time state tables of both the agent's and the subject's expected values for stag and hare hunting for each board configuration, using them to simulate a game prior to making each move in order to select the prey with perceived higher profitability based the simulation's outcome. Instead of simply updating the tables after the agent moves, the agent kept track of a separate set of tables (i.e. Projected state table (subject and agent)) that were updated after the subject's moves in order to simulate the subject's knowledge. This agent was able to learn in real time, using both its own and the other player's behavior, plan ahead based on projected game outcomes, and form a conception of the other player's mental state based on their payoff history. In addition to being used to update the in-game weights after a move on the game board, the actor-critic model was also used between moves to simulate a full game using the player's state tables to approximate the player's likely strategy based on the current setup of the board. Two simulations were conducted, one for each prey, either for 3 moves or until a prey was caught (i.e. recurse > 3 || Temp Stag/Hare == 1), and the prey used in the simulation that ended in the highest Actor state value (i.e. the value in the Actor state table for the ending board configuration) for the agent was selected to pursue during the agent's next move.

The Plan agent used the actor-critic model, but with state tables only for the agent, meaning that the values used to update the state tables did not include the player's position (Figure 3.2C). Like the ToMPlan agent, the Plan agent simulated the game's outcome using the state tables prior to making moves, but without the ability to directly observe the subject's behavior. To achieve this, the state tables were modified so that they did not index by the position of the subject's avatar on the game board. During planning, the player's prey was set to the same prey as the agent, simulating an inability to use ToM to recognize that another individual can hold differing goals and mental states from one's own.

3.1.4. Experimental Design.

A maximum of 40 subjects participated in the experiment simultaneously at each collection session. The computer software portion of the experiment was administered via desktop PCs located in the Experimental Social Science Laboratory. NetSupport School (NetSupport, Ltd) software launched the program and collected the behavioral data in text files. Subjects read written instructions outlining how to play the Stag Hunt, then played a 15-game training block against a random agent. Following the practice, subjects participated in five blocks against different agents (in randomized order), with 50 games per block. Subjects were allowed to complete the task at their own pace. Between each block, the subjects filled out a written survey querying their own behavior and the perceived cooperation of the agent they were playing with.

After finishing the experiment, subjects completed a written demographic survey collecting personal information regarding their age, major, experience with video games, whether English was their first language, and whether they were willing to be contacted for a follow-up study. They then completed the 50-question Autism Spectrum Quotient (AQ) test for adults (University of Cambridge's Autism Research Centre). Subjects were compensated US$7 plus a variable US$0-$15 of performance earnings that corresponded to their score during a randomly selected block of the computer experiment. Each point was worth 7.5 cents, and the total payment was rounded to the closest dollar.

3.1.5 Analysis.

Unless otherwise specified, all analyses were performed in MATLAB (Mathworks, Inc.) and tested for significance using ANOVA and a two-sample Kolmogorov-Smirnov hypothesis test. Each significance threshold was corrected for error of $p < .05$ using Bonferroni correction.

3.1.5.1. Cooperation.

The primary metric of cooperation used in our analysis of the Stag Hunt was the number of successful stag captures per block. In a given block, there was the opportunity to catch 50 stags.

3.1.5.2. Intent.

Human subjects could signal their goal to induce cooperation or intent by their position on the game board. Path analyses included the amount of time a player spent next to the stag attempting to initiate cooperation, and deviation from the optimal path between the subject's start and end game positions. The amount of time spent next to a stag, referred to as "loitering," was measured in seconds from the moment the player reached a square adjacent to the stag until either the player moved to a non-adjacent square, or the end of the game was reached. Path deviation was calculated as the difference between the number of moves taken by the subject and the number of squares between the subject's endgame position and their randomized starting position.

3.1.5.3. Agent State Tables.

The state tables produced by the adaptive agents relays detailed information regarding target choice preferences in relation to game board configuration. State tables were analyzed by counting the number of board configuration states in which the likelihood of selecting a particular target reached or exceeded 60%, indicating preference for that target.

3.2. Results

A number of analyses were conducted assessing various subject factors on cooperation or AQ. We found no significant effects of self-reported English second language status, gender, and previous experience with video games on the number of stags caught and no correlation with AQ. There were also no significant effects of block order on cooperation.

A) B)

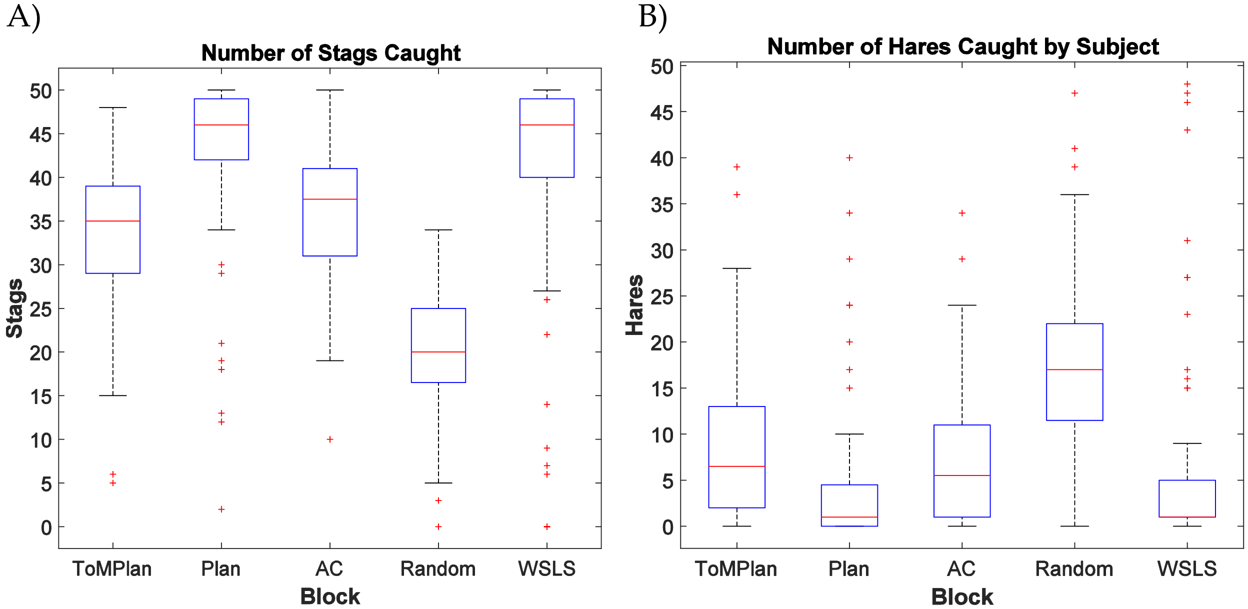


Figure 3.3. Bar graphs for A) the number of stags caught over all games by block, and B) the number of hares caught by subjects over all games by block.
Subjects caught significantly more stags in the Plan and WSLS conditions than all other conditions, and significantly less stags in the Random condition than all other conditions.
Subjects caught significantly more hares in the Random condition than all other conditions, and significantly less hares in the Plan and WSLS conditions than all other conditions. In the graphs,

the x-axis is blocks by agent type, while the y-axis is A) total stags caught during block and B) total hares caught by subject during block, out of a maximum of 50.

3.2.1. Analysis by Agent Type.

Analysis of cooperative behavior (i.e., hunting stags) indicated that subjects were sensitive to the differences between the simulated agents and changed their play accordingly. Subjects caught significantly different numbers of stags and hares depending on the type of agent with which they played. Comparisons over all subjects for number of stags caught between each agent type were significant ($F(1, 4) = 109.49$; $p < 0.001$) in all pairs except ToMPlan v. AC and Plan v. WSLS (Figure 3.3A).

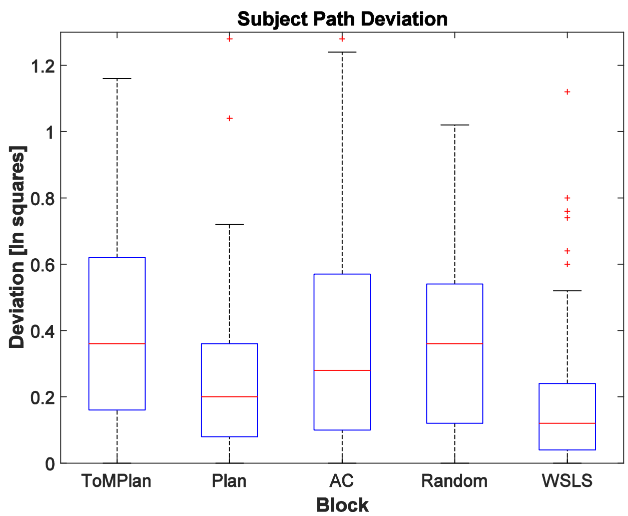
Table 3.2. Behavioral metrics for subjects (mean and standard deviation).

Metric	**ToMPlan**	**Plan**	**AC**	**Random**	**WSLS**
Stags caught	33.365 (8.590)	43.212 (9.081)	35.856 (7.055)	20.019 (6.348)	41.452 (12.385)
Hares caught	8.788 (8.487)	3.567 (7.145)	6.961 (7.043)	17.269 (9.371)	5.212 (10.060)
Subject path dev.	0.412 (0.342)	0.243 (0.236)	0.373 (0.351)	0.377 (0.340)	0.188 (0.230)
Subject loiter	3.754 (0.871)	4.026 (0.719)	3.972 (0.806)	2.551 (0.830)	3.621 (0.935)

Subjects caught the most stags when playing with the WSLS agent and the Plan agent, and the fewest when playing with the Random agent (Table 3.2 and Figure 3.3). Success catching stags with the WSLS and Plan agent indicates subjects engaged in strategies that elicited high levels of cooperation from the agent. For the WSLS, the agent could easily be trained into stag equilibrium after a single trial of successful stag hunting. For the Plan agent, ignorance of the subject's position prevented the agent from being deterred from the stag while the subject was far from the stag target. This led to initial shows of cooperative intent on behalf of the agent, guiding the subject into profitable cooperative behavior. In contrast, subjects had no control over the Random agent, and thus tended to defect when it appeared advantageous.

Similar to stag hunting, an analysis of non-cooperative behavior (i.e., hunting hares) revealed differences in strategy depending on the agent. A one-way ANOVA revealed significant differences in the number of hares caught in games with each agent ($F(1, 4) = 41.12$; $p < 0.001$) between all agent types other than ToMPlan and AC (Figure 3.3B). Subjects caught the most hares when playing with the Random agent, the condition in which consistent cooperation could not be achieved (Table 3.2). This was followed by the ToMPlan and AC conditions, two conditions in which the adaptive agents required multiple game interactions to develop strategies that played effectively with a given subject. The fewest hares were caught in the Plan and WSLS conditions.

A) B)



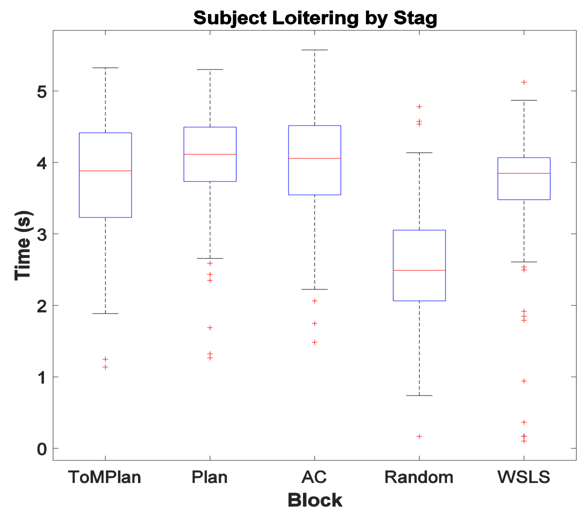


Figure 3.4. Bar graphs of A) average subject deviation from the optimal path in squares during a block and B) average time in seconds spent next to a stag during a block.
Path deviation was significantly less in the Plan and WSLS conditions compared to the ToMPlan condition. Loitering was significantly less for the Random condition compared to all other conditions, which were comparable to each other. In the graphs, the x-axis is blocks by agent type, while the y-axis is deviation in game board squares and time in seconds, respectively.

Subjects also engaged in different visual displays of intent when playing with agents; deviation from the optimal path ($F(1, 4) = 9.52$; $p < .001$) and loitering ($F(1, 4) = 53.79$; $p < .001$) both depended on agent type. Subjects opted for more indirect paths with the ToMPlan agent than the Plan and WSLS agents (ToMPlan v. Plan: $p < .002$; ToMPlan v. WSLS: $p < .001$) (Figure 3.4A; Table 3.2). WSLS and Plan featured the lowest path deviation, meaning subjects typically chose a direct route to their end position during games. Subjects loitered significantly less in the Random condition ($p < .001$), but all other conditions were comparable (Figure 3.4B).

Players used clearly differentiated hunting strategies with each agent type. In the following section, we consider the relationship between subject behavior and autistic traits and we focus on the ToMPlan and Random agent. Because Plan and WSLS conditions largely expressed ceiling effects with the number of stags caught, it was not considered. The AC condition was very similar to the ToMPlan condition, but not as polarizing. Therefore, going forward, we will analyze subject responses to the ToMPlan and Random agents.

3.2.2. Relationship to AQ.

Scores on the AQ test across the subject population ranged from 9 to 36 points, with greater scores indicating higher affectedness by autistic traits. Subjects differed in their behavior and in their responses to different agents depending on their level of autistic traits as measured by AQ. Overall, subjects with high AQ scores tended to cooperate less and had trouble discerning the intent of adaptive agents.

A) B)

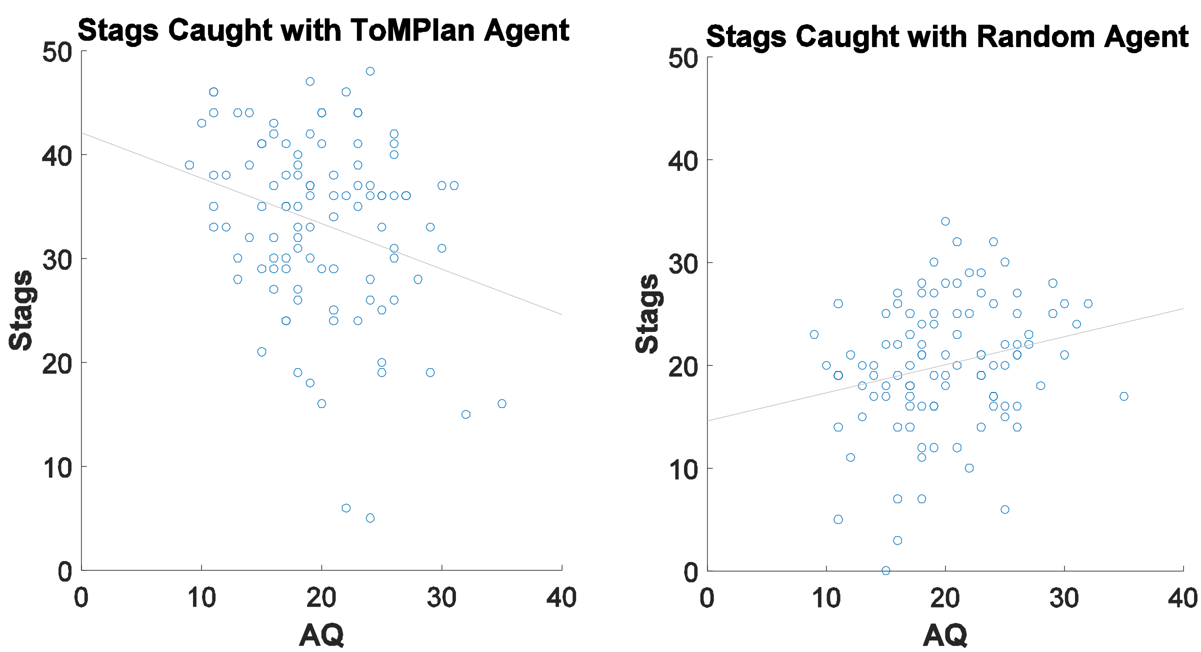


Figure 3.5. Scatter plots of the number of stags caught as a function of AQ when playing with the A) ToMPlan and B) Random agents.
The number of stags caught in the ToMPlan and AC conditions was significantly negatively correlated with AQ, and the number of stags caught in the Random condition was significantly

positively correlated with AQ. The x-axis is AQ ranging from 0-40, the y-axis is the number of stags out of a possible 50, and the regression line is included.

3.2.2.1. Cooperation.

The relationship between AQ and cooperation revealed two distinct patterns depending on agent type. Cooperation, as measured by stag captures, was negatively correlated with AQ in the ToMPlan ($r = -0.276$, $p < 0.002$), and positively correlated with AQ in the Random condition ($r = 0.234$, $p < 0.008$; Figure 3.5 and Table 3.2). Low AQ subjects captured more stags than high AQ subjects when playing with the adaptive agent, whereas high AQ subjects captured more stags than low AQ subjects in the Random condition. The implication is that subjects with high AQ did not recognize and develop strategies for eliciting cooperative behavior from the adaptive agent, evidence for difficulty understanding the adaptive agents' intention.

This does not imply, however, that high AQ subjects did not understand the purpose and goals of the Stag Hunt game. High AQ subjects could cooperate with the simple Random agent, and were even more successful at capturing stags than lower AQ individuals. In this condition, being predisposed to strategies that do not rely on cooperative signals resulted in more optimal play.

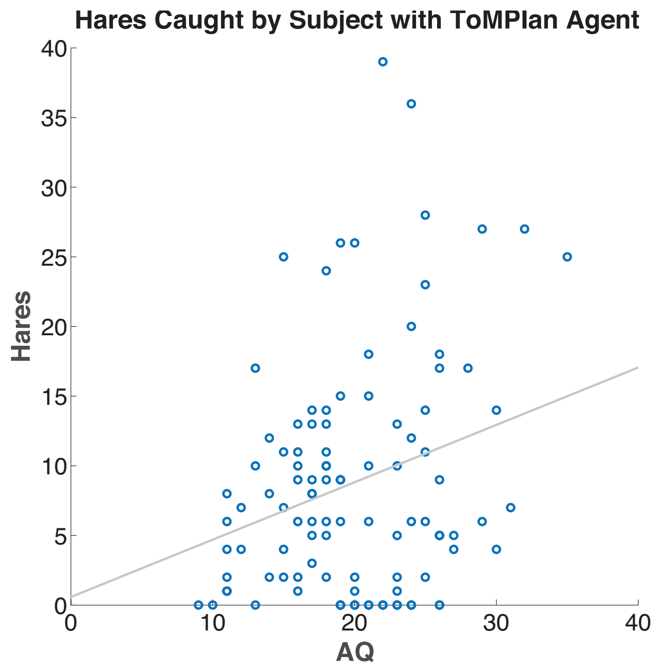


Figure 3.6. Scatter plots of the average number of hares caught by subjects as a function of AQ during the ToMPlan condition.
A significant positive correlation was found in the ToMPlan condition, indicating that subjects with higher AQ tended to catch more hares than in lower AQ subjects. The x-axis is AQ ranging from 0-40, the y-axis is the number of hares out of a possible 50, and the regression line is included.

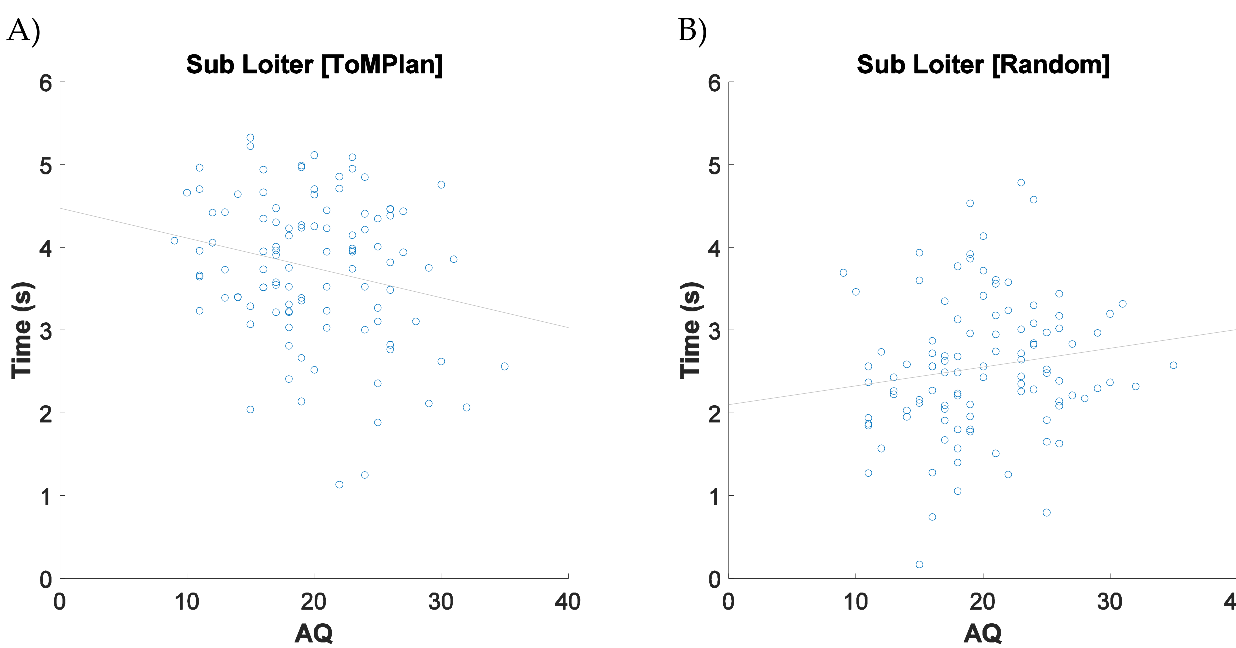
As hare hunting is typically preferential when cooperation cannot be achieved, and subjects with High AQ had difficulty cooperating, there was a positive correlation between AQ and the number of hares caught was observed in the ToMPlan condition (see Figure 3.6A; $r = 0.264$, $p < 0.003$). This significant positive correlation with AQ suggests that high AQ subjects defected more under these conditions because they could not interpret the intent of ToMPlan agent (Schuwerk et al., 2014).

Together, these findings support the hypothesis that high AQ subjects were less likely to engage in implicit exhibit ToM, reminiscent of impairment characteristic of clinical ASD (Begeer, Bernstein, van Wijhe, Scheeren, & Koot, 2012; Senju, Southgate, White, &

Frith, 2009; White et al., 2014). Individuals with high AQ were less cooperative when faced with a player that had the ability to alter its decisions depending on changing context. In other words, they had trouble discerning the intentions of the agent, which simulated attributes of ToM.

3.2.2.2. Intent.

In an effort to understand the connection between AQ and the extent to which subjects attempted to communicate with and influence the other player, we assessed the degree to which subjects signaled their intentions through path deviation and loitering next to a stag during game play. Path deviation is a means by which we can discern whether a subject either used movement to communicate or changed targets during a game, and loitering is an indicator of cooperative intent. Although the analysis of path deviation did not show significant results, analysis of loitering exposed differences between agent types in both subject and agent behavior.



C) D)

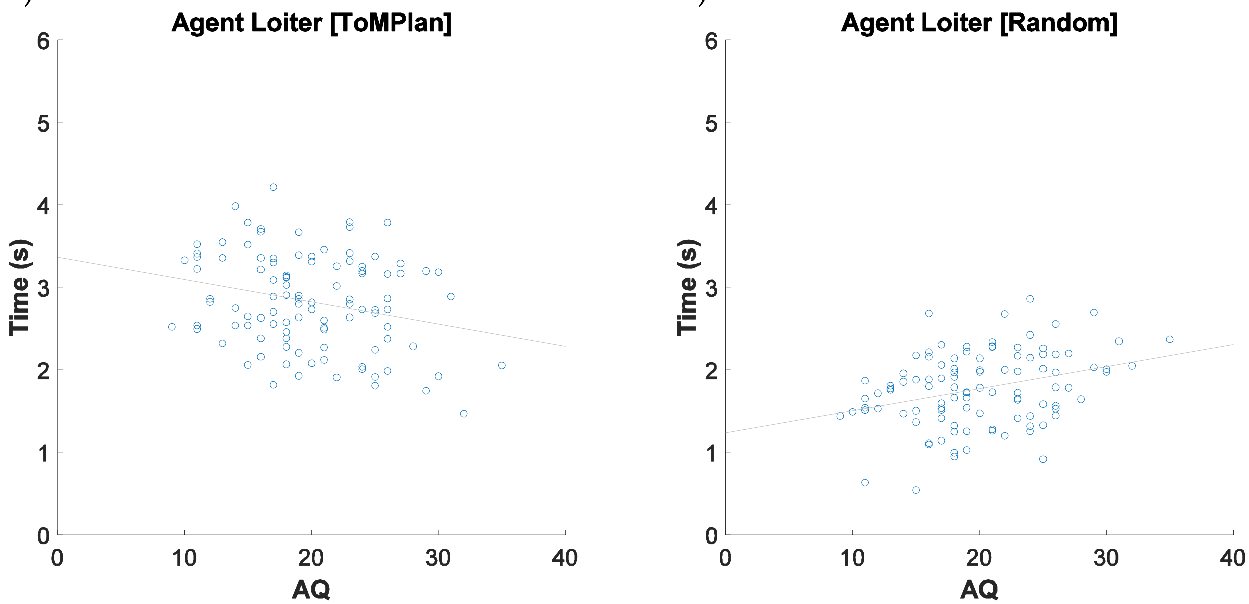
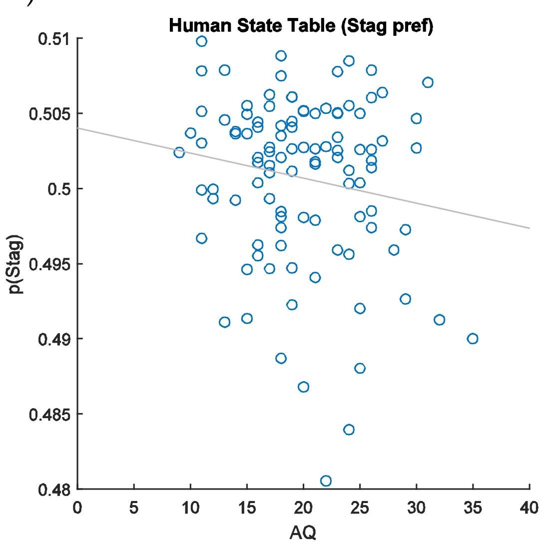


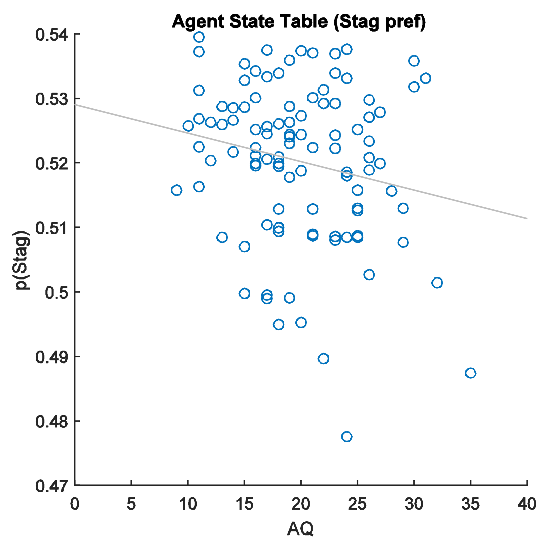
Figure 3.7. Scatter plots of loitering next to stag as a function of AQ.
Loitering was significantly negatively correlated with AQ in games with the ToMPlan agent, both for the subject and the agent (A and C). Loitering was positively correlated with AQ when playing with the Random agent (B and D), although only agent loitering was significant. The x-axis is AQ ranging from 0-40, the y-axis is the average amount of time in seconds the player spent next to a stag per game, and the regression line is included.

The amount of time both subjects and agents spent loitering next to a stag (Figure 3.7) in the ToMPlan condition was negatively correlated with AQ (Subjects: $r = -0.225$, $p < 0.01$; Agent: $r = -0.255$, $p < 0.004$). The amount of time the agent spent loitering next to a stag in the Random condition was positively correlated with AQ ($r = -0.326$, $p < 0.001$), a trend that was seen but did not reach significance in subjects (Figure 3.7B). These effects match the correlation between AQ and stag captures, providing further evidence that subjects with higher AQ tend to spend less time signaling catch intent in the ToMPlan condition, but allow the Random agent more time to signal its intention to catch stags. These trends are potentially indicative of an increase in systemizing behavior on behalf of the high AQ subjects, as the Random agent was predictable within game due to its direct path toward its target and therefore, high systemizing subjects could pick up on the pattern and use it to their advantage. The ToMPlan agent on the other hand incorporated unpredictability both within and between game as a result of its constantly adapting behavior, making it more difficult to recognize a pattern to systematically respond to.

A)

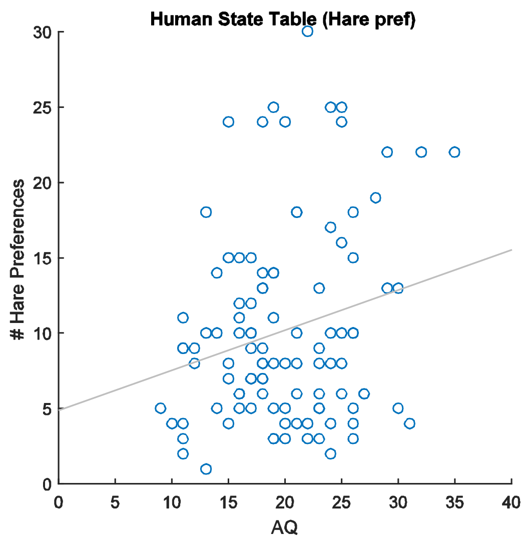


B)



C)

D)



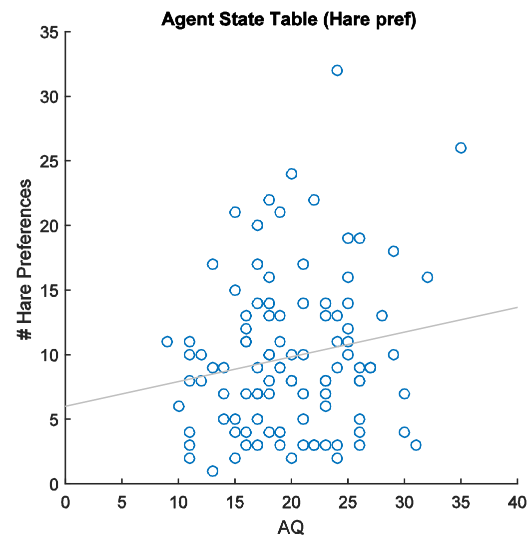


Figure 3.8. Scatter plot of target preference according to human and agent state tables collected by the ToMPlan agent.
A) and B) show the average state table weights for stag preference for the human and agent, respectively. In C) and D), the number of target preferences for a subject was defined as the number of states in each table in which the likelihood to hunt the target was greater than 60%. There were significant positive correlations between subject AQ and the propensity for the ToMPlan agent to hunt stag, and the number of hare preferences and the subject's AQ score. The x-axis is AQ ranging from 0-40, the y-axis is the average state table weight in propensity to hunt stag (A & B) or the number of target preferences (C & D), and the regression line is included.

3.2.2.3. Agent State Tables.

In analyzing the state tables created by the ToMPlan agent, we were able to find trends that underscore the findings above (Figure 3.8). Subject AQ and the state table weights were negatively correlated, indicative of the agent's propensity to hunt stag when playing with low AQ subjects ($r = -0.195$, $p < .05$; Figure 3.8B). The same trend was found in the subject state tables, although this did not reach significance ($r = -0.16$, $p < 0.1$; Figure 3.8A). Analysis of the ToMPlan's human state tables derived from subject choice preferences and path information revealed that the amount of states in which subjects showed a preference for hare positively correlated with AQ score ($r = .224$, $p <$

.022; Figure 3.8C). The same trend was found in the agent state tables, although this did not reach significance (r – 0.17, $p < .08$; Figure 3.8D). This indicates that subjects with higher AQ showed they were less amenable to stag hunting in a way that was apparent to the ToMPlan agent.

3.3. Discussion

Results of the present study successfully showed that strategizing and decision-making in social game environments is differentially linked to the level of autistic traits in subjects of the general population. We found differences in the tendency to elicit and engage in cooperative stag hunting, and in the analysis of intent through subject movements across the game board.

Subjects with higher incidence of autistic traits failed to engage in social behaviors when playing games with agents simulating ToM and planning. AQ score was negatively correlated with the amount of cooperation when playing with an agent that simulated ToM, and AQ score positively correlated to the amount of cooperation when playing with a simplistic, randomly acting agent. This suggests that subjects with high levels of autistic traits may have had trouble understanding the intent of an agent that simulates ToM, much like individuals with clinical ASD diagnoses (Boucher, 2012; Kana et al., 2009; Mason et al., 2008; Matthews et al., 2012; Senju, 2012).

The finding that individuals with high AQ scores do not have as much difficulty cooperating with Random agents as low AQ individuals suggests that when higher AQ subjects can discern another's strategy or intention, they might be more capable of cooperation. This theory is also seen in the positive correlation between time spent loitering next to the stag and AQ, as evidenced by the increase in visual intention signaling by the Random agent. The negative correlation between AQ and loitering for the ToMPlan agent in both subject and agent shows the converse, as subjects with high AQ likely spent less time signaling cooperative intent, suggesting that subjects with higher levels of autistic traits may have had more trouble discerning the strategies of this agent, indicative of underlying social deficits (Simon Baron-Cohen et al., 1999; Feil-Seifer & Matarić, 2009; Rolison, Naples, & McPartland, 2015). Taken together, these results demonstrate that a non-verbal task, such as the Stag-Hunt, can induce differential behavior among subjects whose autistic traits differ, and that these differences may be related to variations in implicit theory of mind.

3.3.1. Agent Differences.

Computer agents that learn from experience have been used in both embedded and embodied implementations to study strategy formation and learning with and without human subjects (Baldassarre, 2003; Kidd & Breazeal, 2004; Merrick, 2010; Merrick & Maher, 2009; Schembri, Mirolli, & Baldassarre, 2007; Sutton, Barto, & Williams, 1992; Szolnoki & Perc, 2009; Szolnoki, Xie, Ye, & Perc, 2013; Thomaz & Breazeal, 2008; Valluri, 2006). Like many of these studies, the present study utilizes a model implementing reinforcement learning, which is practiced in biological systems (Frank et al., 2015; Gläscher, Daw, Dayan, & O'Doherty, 2010; Glimcher, 2011). The actor-critic model (Khamassi, Lachèze, Girard, Berthoz, & Guillot, 2005; Li, Lowe, & Ziemke, 2013; Nakamura, Mori, Sato, & Ishii, 2007) formed the foundation for the three adaptive agents in this study, which enabled them to learn in a real-time environment, tailoring their gameplay to the unique strategies of the subject with which they were currently engaged.

While the AC agent used the base actor-critic model and was mostly reactive, planning only one action in advance and considering the position of the other player but not their likely actions, the ToMPlan agent added both planning and mental simulation to consider the future and the likely actions the subject would take when determining a prey to pursue. Both agents began a block of games without bias and acquired behavioral tendencies through their turn-by-turn experience with a given human player. These agents were designed for the present study in order to present a complex social partner to subjects, similar to a human player but lacking the inherent bias and confounds when using human social interaction. While the AC agent was shown to be sufficient in evoking complex and varied reactions from subjects in our past study (A B Craig et al., 2013), the new elements of simulated ToM and planning made the ToMPlan agent more polarizing in terms of subject reaction. In other words, the ToMPlan and AC agents largely evoked the same responses in subjects, but the ToMPlan's effect was stronger.

Although the Plan agent used similar underlying architecture to the ToMPlan and AC agents, subject behavior when playing with this agent was most comparable to play with the WSLS agent; both agents elicited a strong cooperative strategy from subjects. The Plan agent was included as a control agent with the ability to simulate games to inform action decisions, but without the ability to model future actions of the human player. Because the Plan agent did not consider the subject's behavior or placement on the game board, it frequently began a block of games attempting to pursue the most profitable prey, the stag. Most subjects ascertained this signal of cooperation, which is likely responsible for the near ceiling numbers of stags caught in the Plan condition.

Like the Plan agent, the WSLS agent also readily elicited cooperation from the human subjects. The WSLS agent (Imhof et al., 2007; M. Nowak & Sigmund, 1993) kept its target following a win and switched its target after a loss. As a result, once the first stag was caught in a block of games, the WSLS would continually pursue stags until the subject defected. It was easy for subjects to fall into a cooperative equilibrium with the WSLS agent; the subjects just needed to move their avatar directly to the stag every time. The WSLS agent was included as a control agent, influenced by subject behavior but with a simple, fixed strategy that did not vary in real time like the adaptive agents (see Table 3.1). However, because the WSLS and Plan agents did reach near ceiling levels of cooperation, games largely became trivial, moving straight to the same target over and over.

Lastly, the Random agent was included in the present study as a baseline control for all other agents, given that it held no strategy, was not affected by subject behavior, and its behavior did not change in any meaningful way over time.

3.3.2. Effect of agent type on behavior.

The ToMPlan agent evoked behavioral responses in subjects that did not reach the magnitude of cooperative equilibrium. Compared to the other agent types, in games with the ToMPlan agent, the ratio of hares to stags caught by subjects was midway between the highest and lowest conditions (Figure 3.3; Table 3.2). Agents in this condition operated off of constantly changing state tables that learned in a way that is not transparent to the subjects. Even if a subject attempted to cooperate or defect consistently with the agent, the state tables learn based on board configurations that can be ambiguous, meaning that it is difficult for the agent to behave consistently in every game. In addition, the agent's response is granular, functioning on the level of individual moves rather than the highly integrated cognitive response that humans produce. The moderate number of stags in the ToMPlan condition may have been a result of the uncertainty of the actor-critic model, which added difficulty to the formation a cooperative equilibrium with the subject.

In games with the Random agent, overall findings show a highly polarizing effect on subject behavior. The Random agent evoked the least amount of stag hunting and the highest number of hares caught by the subject out of all of the agent types (Figure 3.3; Table 3.2). Unlike the adaptive agents, the Random agent could not be coerced into cooperation under any circumstance. As the number of stags caught in games with the Random agent was close to half of the maximum number of stags available in a block, this is in keeping with the rules of probability that govern the Random agent. Once the subject ascertained that their behavior could not influence the agent, their action decisions became relatively straightforward. Similar decreases of engagement when playing with a randomly-acting agent have been seen in simple games, both behaviorally and neurally (Chaminade et al., 2012). When considering the paths themselves, path deviation did not significantly differ between the ToMPlan and Random conditions (Figure 3.4A). Although this finding is seemingly at odds with our previous finding that subject take longer paths when playing with complex adaptive agents (A B Craig et al., 2013), it is possible that for some subjects, the probabilistic nature of the Random agent in fact makes that agent complex and increases the amount of uncertainty in interacting with it, leaving subjects desiring more information to deal with this type of behavior (Hilbert, 2012; Schultz, Mitchell, Harper, & Bridges, 2010).

3.3.3. Effect of AQ on behavior

3.3.3.1. Cooperation.

Subjects with high AQ had a difficult time deciphering the intentions and influencing the strategies of the agent. Specifically, we observed a significant negative correlation of AQ to the number of stags caught in the ToMPlan condition (Figure 3.5A), which was accompanied by a significant positive correlation of AQ to the number of hares caught in the same conditions (Figure 3.6). Because the ToMPlan agent modeled aspects of ToM, understanding this agent required a higher level of ToM processing. With this in mind, the fact that high AQ subjects did not seem to be as successful as low AQ subjects at reaching a cooperative equilibrium with the ToMPlan agent indicates that the high AQ subjects failed to engage implicit ToM when implementing their game playing strategy. This behavior of individuals with high AQ is reminiscent to that of individuals with clinical ASD (Kana et al., 2009; Matthews et al., 2012; Senju, 2012).

Interestingly, subjects with high AQ were more successful than subjects with low AQ at catching stags with the Random agent. Positive correlations with AQ score and the Random agent were found in the metrics of stags caught (Figure 3.5B). This result eliminates the possibility that high AQ subjects were simply less successful in cooperation when playing the Stag Hunt in general. One potential explanation for this difference could be that low AQ subjects, after several failed attempts at cooperation, became upset with the agent and acted out against them by defecting. Previous studies in game theory have shown that emotion plays an important role in decision-making during social games (Rilling & Sanfey, 2011; Sanfey, 2003; Schultz et al., 2010; Schwarz, 2000). The phenomenon of acting against one's best interests in order to seek revenge on a party that one believes has wronged them is well established in game theory, especially in tasks such as the Ultimatum Game (M. A. Nowak et al., 2000; Rand, Tarnita, Ohtsuki, & Nowak, 2013; Wei, Zhao, & Zheng, 2013; Wout, Kahn, Sanfey, & Aleman, 2006). In the case that high AQ subjects did not consider the other player's mental state, this emotional valence may not have been strong enough to desire revenge or to trigger negative emotional response to the agent's behavior at all. As empathy and emotion is interlinked with ToM processing (Boucher, 2012; Ciaramelli, Bernardi, & Moscovitch, 2013), the difference between high and low AQ subjects likely revolves around the recruitment of ToM in decision-making. Alternatively, high AQ subjects could have been more successful with the Random agent because of an overlying difficulty understanding the broader strategies of agents. This limitation would suggest that high AQ subjects would need to strategize on the level of games rather than on the level of blocks, a tactic that would allow for capitalizing on the trials in which the Random agent happened to go for a stag rather than discontinuing attempts at cooperation with an agent who could not consistently cooperate.

3.3.3.2. Strategizing under uncertainty.

In games with the ToMPlan agent, AQ score affected the amount of time players loitered next to a stag. With the Random agent, AQ positively correlated with both loitering time for both the subject and agent (Figure 3.7B & D), but negatively correlated with the ToMPlan agent (Figure 3.7A & C). It stands to reason that subjects with higher AQ spent longer in games with the Random agent in an either in an attempt to gain more information from the agent to determine its strategy, or because the subject did not understand the agent's goal at the beginning of the game and changed his/her prey choice midgame. Both of these conclusions are indication of the overarching theme that high AQ subjects have more difficulty understanding the behavior of adaptive agents, which likely stems from implicit ToM impairment (White et al., 2009).

Analysis of the state tables collected by the ToMPlan agent showed that the agent's preference for stag hunting decreased in games with subjects of high AQs (Figure 3.8B). This correlation is evidence that the agent was able to pick up on the subject's behavioral cues and shape its own behavior to match that style of play. While not significant, the human state tables displayed the same trend of negative correlation to AQ score (Figure 3.8A). In regard to hare hunting, the state tables revealed that subjects with higher AQs preferred hare hunting with this agent (Figure 3.8C), a trend also found in the raw number of hares and stags caught. Although not significant, this same preference for hare hunting was found in the agent state tables (Figure 3.8D). Coupled with the results on loitering, it is clear that high AQ subjects had more trouble understanding the ToMPlan agent's cooperative intent and therefore spent less time signaling visual intent. Furthermore, the subject state tables illustrate the agent's ability to accurately assess the target preferences of the subjects. The ToMPlan agent learned that high AQ subjects are less likely to hunt stags, and adjusted its behavior accordingly. While reinforcing the theories above, this information is important for its insight into what the model controlling the ToMPlan agent is perceiving. The values collected in the agent's state tables allow us to predict the subject's behavior, and this predictive value allows us to estimate the AQ of the subject. Simply through the information the agent acquired during previous games, the agent learned that subjects with higher AQ were less amenable to cooperation. Through the state table values, it is apparent that the subject's behavior shaped the agent's behavior towards hare hunting in games with high AQ subjects.

When assessing these findings with the results of cooperation, in the ToMPlan condition, low AQ subjects tended to invest a longer amount of time signaling the intention to hunt stags, but when they went for a stag, they tended to be successful. When high AQ subjects were able to catch a stag, they tended not to have loitered as long as the low AQ subjects to do so. It is possible that the positive correlation with AQ and hare outcomes resulted from high AQ subjects being unable to induce cooperation and switching to hare hunting during a game. Because the hare outcome can be attained without any path deviation or extra movement, it is likely that subjects who take many moves during a game in which they catch a hare did not begin the game with a commitment to hare hunting. This explanation would further affirm the notion that high AQ subjects have difficulty deciphering the ToMPlan agent, as it implies they were either unable to get the ToMPlan agent to cooperate or did not understand how to influence ToMPlan agent.

In the Random condition, on the other hand, when playing with high AQ subjects, the agent tended to spend longer loitering next to the stag, but were more often successful than with low AQ subjects. Because the Random agent operates on a set strategy, this difference resides in subject behavior. High AQ subjects spent more time observing in games with the Random agent than low AQ subjects, which caused the Random agent, who would move directly to a target and wait, to spend more time waiting next to a stag. This explanation implies that subjects with higher AQs invested more time in determining the Random agent's intent visually before deciding what target to pursue, whereas subjects with lower AQs chose a target quicker and did not spend as much time observing the Random agent's behavior. The fact that in the present study high AQ subjects seem to cooperate better with a Random agent is in keeping with the high systemizing mentality, as it was likely easier for these subjects to analyze the patterns of the Random agent. The low AQ subjects, relying more on their intuition as to the mental state of the agent, would not receive useful cues, as the Random agent behaved without a unified goal between games.

3.3.4. Additions to Prior Research.

The concept of explicit vs. implicit ToM tasks is a key component in using ToM tasks to study autism, as it has been theorized that the atypical activity of the ToM network in the brain differs depending on which type of task is employed (S. J. White et al., 2014). It is especially important to investigate implicit ToM, as the process arises spontaneously and is therefore more difficult to accurately assess. The present study adds to the comparatively small field of implicit ToM research, creating a non-verbal paradigm to evoke implicit ToM through cooperation. We have shown through metrics of cooperation and intention signaling that autistic traits correlate to differential behavior in this task, likely with the difference residing in ToM impairment.

In the Yoshida et al. studies (2008; 2010a; 2010b), agents lacked the ability to react in real-time, instead switching their level of sophistication at random. This behavior does not lend itself well to the assessment of social interaction, as the other player would not reliably behave in a manner that appeared to be influenced by the subject. This study goes beyond Yoshida's work in that the agents used in this experiment adapt in real-time rather than switching between fixed models at a designated interval. The adaptive agent entered the Stag Hunt environment naïve and developed a strategy over repeated games based on both the behavior of the other player and the outcomes of the games. We were able to show that subjects utilized more complex paths when playing games with an adaptive simulated agent that developed strategies over repeated games based on both the behavior of the other player and the outcomes of the games. This was in comparison to play with fixed strategy agents, which implemented the same hunt strategy regardless of the human path or game outcomes. With the present study, we built upon the prior work to investigate the between-subject differences as related to levels of autistic traits by using computer agents with behaviors that range from fixed strategy to simulated ToM to probe subjects' ToM abilities. Utilizing various types of adaptive agents in this experiment allowed more nuanced observation of behaviors in subjects. Metrics such as loitering next to a stag were heavily influenced by the ability of the agent to change its prey choice within a game, providing more detail into the formation of strategy and the utilization of paths as a signifier of intention.

3.3.5. Limitations.

One of the surprising results of the present study was we had initially attempted comparing the twenty highest and lowest AQ subjects on the metrics used above, but the findings failed to reach significance. A possible explanation for the inability to find more significant results between the High and Low groups could be the fact that the subject population did not possess high enough AQ scores to show substantial differences from our Low AQ group. We performed all of the previously discussed analyses with twenty of each of the highest and lowest scored AQ subjects, but the subject population may not have possessed a sufficiently wide range of AQ scores. A recently conducted meta-analysis of studies conducted using both neurotypical and clinically diagnosed autistic individuals found that their clinical population averaged 35 points on the AQ test (Ruzich et al., 2015). Only a handful of subjects in the present study scored 28 or above, with the mean score for the High AQ group equaling 27.75. In this case, the size of the subject population was a limitation that could be improved upon by expanding the number of subjects until the Low and High AQ groups featured more extreme scores.

Another possible explanation for the lack of significant results between High and Low AQ groups is that because our subject population was comprised of adult college students, the presumption is that these individuals are highly functional. The implication is that differences resulting from affectedness by autistic traits may only be observable through neural processing because behavior has been corrected to achieve high functioning status through neural compensation. In this case, an investigation into the brain functioning occurring during gameplay would be necessary to display differences between the Low and High AQ groups. Such a continuation on the present study would also afford comparisons to other studies in the growing body of research focused on neuroimaging correlating AQ among the general population with irregular neural processing attributed to clinically autistic individuals.

The findings of the present study indicate a strong connection between the ability to cooperate with adaptive and fixed strategy agents and the individual's level of affectedness by autistic traits as determined by AQ score. Likely as a result of a decrease in consideration of other individuals' mental states, subjects with higher AQs tend to cooperate less with adaptive agents than subjects with lower AQs while cooperating more with agents that behave randomly. Although this finding is apparent in select behavioral metrics, there is likely a strong underlying cognitive component to this difference that can only be observed through the addition of information regarding neural processing. In future work, we hope to utilize fMRI to investigate differences within the ToM network of the brain between low and high AQ subjects playing the Stag Hunt with these computer agents. This work will help to further the field's understanding of autism spectrum disorders as a gradient within the general population by identifying atypical neural correlates that are common between individuals with formal ASD diagnoses and neurotypical individuals with high AQ. The Stag Hunt software developed for this study holds potential for use in clinical autistic populations as a non-verbal game capable of probing ToM abilities without the confound of human interaction.

CHAPTER 4: Neural Basis of Theory of Mind Processing Differences Between Subjects of Varying Levels of Autistic Traits

The present body of work identified behavioral components related to autistic traits, theory of mind, and cooperation. In addition, it would be of interest to identify the neural correlates of this behavior. Therefore, the next addition to this line of research is the utilization of functional magnetic resonance imaging (fMRI) to identify the potential differences in neural activation both between subjects of varying levels of autistic traits and when playing the Stag Hunt game with agents of varying strategies.

In this preliminary study, we used fMRI in conjunction with adaptive agents while playing the Stag Hunt game, combining a ToM implicit, game theoretic task environment with intelligent and autonomous yet artificial agents in order to analyze the neural and behavioral response to non-human agents that probe ToM.

4.1. Specific Aims & Hypotheses.

The specific aims of the proposed research are as follows: 1) Utilize a computer agent incorporating functionality based on aspects of ToM and is capable of evoking ToM behavior, in subjects during the Stag Hunt. To make decisions, this agent takes into consideration (i) its own potential costs and rewards, (ii) the other players potential costs and reward, and (iii) recursive planning of potential outcomes. 2) Investigate differential behavioral responses between low AQ and high AQ subjects towards adaptive agents controlled by ToM model when playing the Stag Hunt. The novel addition of adaptive agents to the experimental paradigm will engage a complex, ToM response in subjects not possible using set-strategy agents and more controlled than human opponents, a quality necessary for assuring the validity of conclusions. We expect to find response patterns that match most closely set-strategy agents in subjects with high levels of autistic traits, indicating a reduced ToM response, while subjects with low levels of autistic traits are expected to match adaptive agents, indicating a higher degree of ToM response. 3) Investigate differential neural response between subjects of varying levels of autistic traits in the mentalizing network through fMRI when playing the Stag Hunt with adaptive agents. We expect to find hypoactivation of ToM network in subjects with high AQ compared to subjects with low AQ in implicit ToM tasks, and hyperactivation in explicit tasks, responses that correlate with behavioral findings in Aim 2) and identify the nature of atypical ToM network activation for subjects with high levels of autistic traits in explicit vs. implicit ToM tasks.

4.2. Experimental Paradigm.

The neurobiologically-inspired model capable of evoking ToM response in subjects is equipped with mentalizing and planning capabilities, key aspects that comprise and cue ToM response (see Chapter 3).

This agent, alongside fixed strategy control agents, will be used as players in an fMRI experiment to test hypotheses regarding the theory of hypo- and hyper-activation of the mentalizing network in subjects with high levels of autistic traits when using implicit vs. explicit ToM tasks. Several behavioral measures will be taken during games including score, prey preference, path trajectory, and game length, which will be compared both between both high and low AQ subjects, as well as to the adaptive agent data as tested above, to make conclusions regarding what mechanisms were likely favored in each group's play style. We expect the results of the proposed behavioral study to show a decrease in cooperative behavior in high AQ subjects compared to low AQ subjects, similar to what has been shown in past research with neurotypical subjects and subjects with clinical ASD (Yoshida, Dziobek, et al., 2010). As cooperative behavior is subserved by the ability to imagine the mental state of another person, the implication of this finding is that subjects with high AQ may engage less in ToM behavior when performing implicit ToM tasks such as the Stag Hunt. We expect the results of the proposed fMRI study to show hypoactivation in the mentalizing network (i.e. mPFC, TPJ, APC) in high AQ subjects vs. low AQ subjects when playing the Stag Hunt game, as it does not explicitly cue ToM behavior to take place and therefore should not strongly activate in subjects with high levels of autistic traits even though subjects with low levels of autistic traits should engage spontaneous ToM. We expect hyperactivation of the mentalizing network in high vs. low AQ subjects when responding to questions, which explicitly cue ToM, a finding that has been observed in past research (White et al., 2014). These two findings would provide evidence in support of the explicit vs. implicit ToM theory, as well as insight into atypical activity and connectivity of the mentalizing network in the high AQ brain.

4.3. Pilot data.

Preliminary efforts have already begun to collect pilot data using general population individuals assessed by AQ score. The preexisting Stag Hunt software was modified for use in the fMRI scanner to accommodate the added restrictions inherent to the testing environment. Nine subjects were acquired through either the Experimental Social Sciences Laboratory or on campus advertisement at the University of California, Irvine. Subjects filled out the Autism Quotient (AQ) test to be used as a metric for evaluating their cooperative performance relative to other subjects of varying AQs. They were selected for participation based on AQ if they scored either lower than 15 or higher than 25. The range for low AQ was 9-14, and the range for high AQ was 26-31. Subjects played repeated games in three, thirty-second blocks separated by 10 second rests of the ToMPlan agent, the Random agent, and the WSLS agent (see Chapter 3) interleaved. Subjects were not told which agent they were playing with, although they were notified when the agent strategy would change. Rather than using the computer keyboard, subjects used two, four-button control boxes for moving their player icon around the game board. During games, fMRI data was collected at two-second intervals to capture neural activity for comparison both between subjects and between agents.

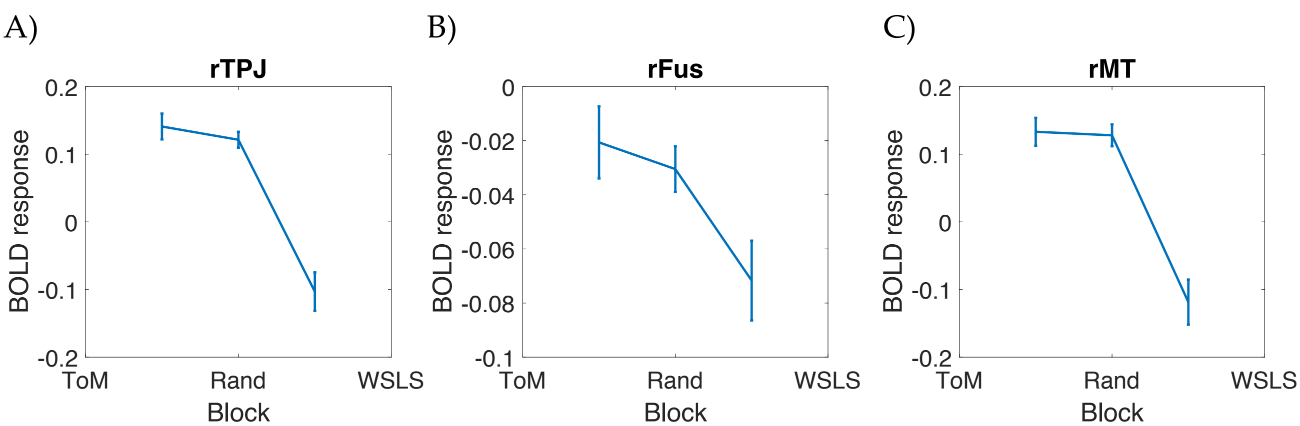


Figure 4.1. Plot of average change in BOLD response during games in each condition. A) Right temporoparietal junction, B) right fusiform gyrus, and C) right medial temporal gyrus each exhibited similar trends indicating higher activation when playing with the ToMPlan and Random agents than the WSLS agent. This is likely due to the fact that the ToMPlan and Random agents required more effort to understand their strategy, whereas the WSLS agent operated on a fixed strategy.

Although we were only able to acquire the data of nine subjects, results of this pilot data reveal trends in BOLD response related to both agent type and level of autistic traits in the subjects. Average change in BOLD activation indicates that the highest level of activation was exhibited in games with the ToMPlan agent, closely followed by the Random agent in rTPJ, right fusiform gyrus (rFG), and right medial temporal gyrus (rMT) (Figure 4.1). Activity in these areas was substantially less in games with the WSLS agent. These trends indicate that subjects expended more effort in games with the ToMPlan and Random agents than with the WSLS agent. The ToMPlan and Random agents are complex in their strategy, given that a subject cannot discern with certainty what the agent will do from one game to the next. The WSLS agent operated on a fixed strategy, so it was possible to predict the agent's behavior on each game. Given that rTPJ has been implicated in ToM processing (Gallagher & Frith, 2003; R Saxe & Kanwisher, 2003), its increased activation in games with agents using complex strategies is unsurprising. In regard to the increase in activity in rMT, this is also potentially related to the role of rMT in theory of mind processing (Sabbagh, 2004), which works in tandem with the rTPJ to contribute to processing of social stimuli. Rather than increased activity, rFG expressed deactivation during games, although the trend is the same as that found in the other brain areas. rFG has been found to play a part in the processing of visual stimuli and in some cases has shown up in theory of mind processing (Gallagher et al., 2000; Gobbini, Koralek, Bryan, Montgomery, & Haxby, 2007; Martin & Weisberg, 2003).

A) B) C)

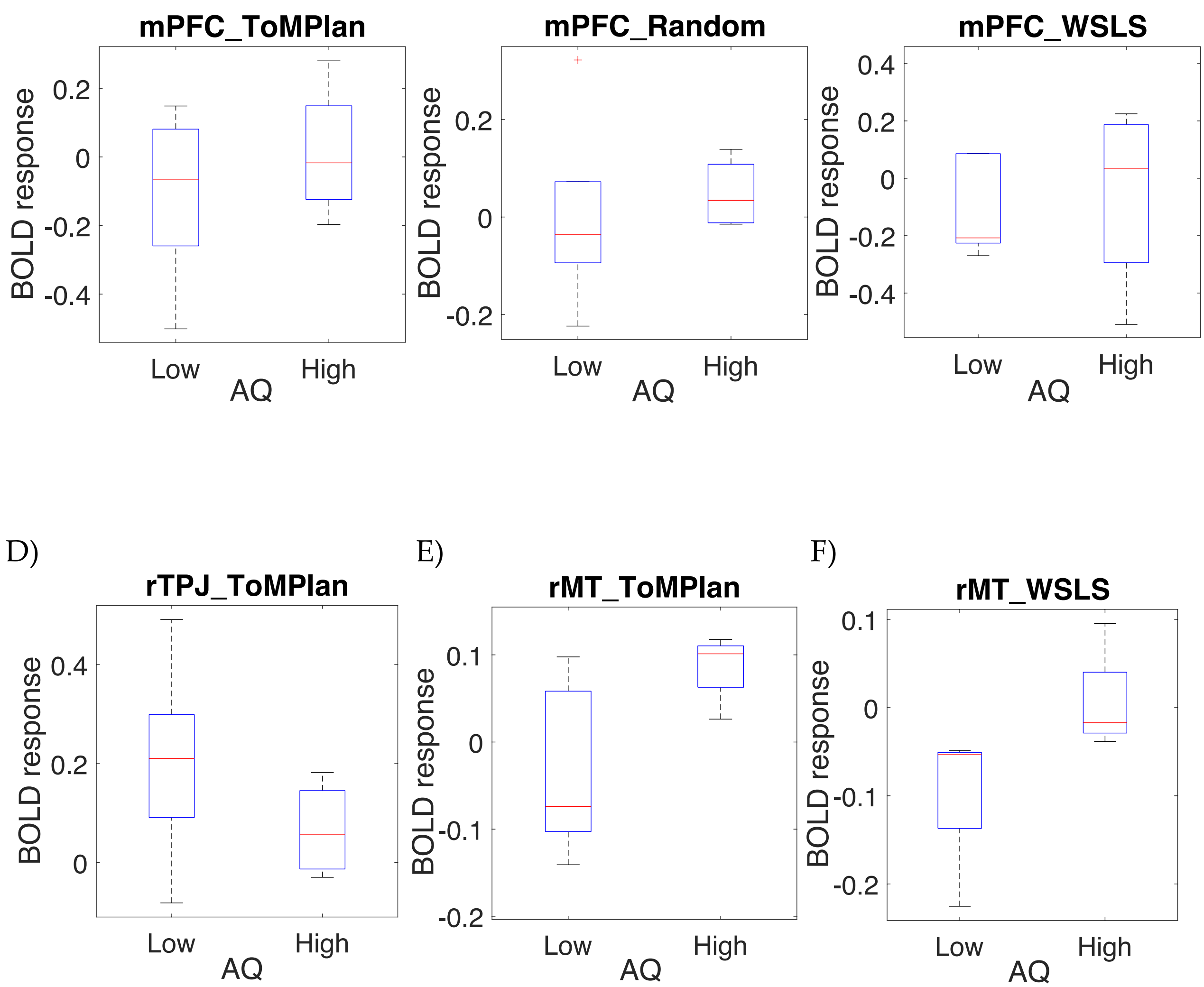


Figure 4.2. Box plots of change in BOLD response between subjects with low and high levels of autistic traits.
A), B), and C) show the BOLD response within the medial prefrontal cortex during games with the ToMPlan, Random, and WSLS agents, respectively. All conditions exhibit trends indicating increased activation in high AQ subjects. Additionally, this trend is found in the right medial temporal gyrus for the E) ToMPlan and F) WSLS agents. The opposite trend is found in right temporoparietal junction for the ToMPlan agent. The wide range of variance resulting from small sample size prevents clear conclusions, but trends indicate low AQ subjects did not use mPFC and rMT strongly in games, especially in the ToMPlan and WSLS conditions. Additionally high AQ subjects recruited rTPJ less in games with the ToMPlan agent.

To analyze BOLD response on the basis of AQ, subjects were split into groups of five low AQ and four high AQ. In mPFC, low AQ subjects tended toward lower BOLD response in all conditions compared to high AQ subjects (Figure 4.2. A-C). The same was true of rMT in games with the ToMPlan and WSLS agents (Figure 4.2 E-F). mPFC is linked to higher level cognition, planning, and decision-making behavior (Barraclough, Conroy, & Lee, 2004; Benoit, Szpunar, & Schacter, 2014; Nyhus & Barceló, 2009), while rMT has been shown to be involved in theory of mind assessment (Sabbagh, 2004). However, the trend found in rMT exhibited deactivation, a decrease in activation as compared to the baseline. In recent studies, deactivation of brain areas has been traced back to the default mode network, which keeps many areas of the brain active in a steady state during rest (Mars et al., 2012). For this reason, some areas of the brain become less active when recruited in tasks such as the present study. This heightened deactivation in low AQ subjects could be a result of these subjects recruiting less higher level processing when playing games, indicating that these subjects had an easier time coping with the agents' uncertain strategies. Lastly, rTPJ was found to be more active for low AQ subjects in games with the ToMPlan agent (Figure 4.2 D). As rTPJ processes ToM judgments (Michael V. Lombardo, Chakrabarti, Bullmore, & Baron-Cohen, 2011, p. -; R Saxe & Kanwisher, 2003), it follows that low AQ subjects dealt with the increase in uncertainty caused by the ToMPlan agent by recruiting rTPJ. High AQ subjects recruited this area to a much lesser degree, indicating a potential deficit in processing of ToM. It is possible that the increased processing in areas like mPFC was an effort to compensate for the deficit in normal social processing areas like rTPJ so that high AQ subjects could process uncertainty, albeit in a more systematic, problem-solving capacity. Such compensatory efforts have been shown in individuals with clinical ASD diagnoses (S. J. White et al., 2014).

Taken with the caveat that the subject pool would ideally increase to 20 per condition, the pilot data collected identifies key trends that are both in keeping with the literature and add validity to hypothesized effects. In future work, I plan to expand upon this study to gain a clearer picture of the neural activity of subclinical autistic traits in the general population. This line of research will help to bridge the gap between ASD research and subclinical social deficits, hopefully with the intent of deciphering the origins of the disorder and the development of its neural expression.

CHAPTER 5: Implications of Findings and Future Directions

The preceding chapters documented a body of work that comments broadly on the methods by which humans cope with uncertainty in their environment to make decisions. Using adaptive agents to evoke complex reactions in subjects without the inherent unpredictability of human social interaction, this work identifies strategies and cognitive biases humans use when engaging in social and economic games that mirror situations that exist in daily life. This work also explores a population that typically struggles with social interaction by investigating individuals in the general population who exhibit high levels of traits characteristic of individuals with ASD, both in behavioral and neural data.

Chapter 1 detailed the use of adaptive agents in a spatiotemporal variant of the Stag Hunt game to assess differences in subject cooperation between agents of differing strategies. Subjects were found to have more variation and uncertainty when playing with the adaptive agent, taking significantly more turns than in games with the Random agent. This finding indicated that players may have sensed the malleability of the adaptive agent, prompting them to try to influence the agent or by changing strategies in-game once the agent's actions became apparent.

In Chapter 2, a probabilistic component was incorporated into the Wisconsin Card Sorting Task to investigate the use of cognitive biases in varying degrees of uncertainty. Subjects increased usage of the risk-averse Observe option during trials of high uncertainty, and tended to overselect the lowest probability rule while underselecting the highest probability rule. These results indicated a tendency for subjects to respond with a decreasing capacity for optimal decision-making behavior under high levels of uncertainty.

In Chapter 3, the spatiotemporal stag hunt game was used as a foundation for further investigation of cooperative behavior while playing with agents of varying strategies, and the level of autistic traits in general population individuals was correlated to performance. The actor-critic model-based adaptive agent was redesigned to incorporate abilities simulating theory of mind and planning to more realistically imitate a human player. Subjects with higher levels of autistic traits, as assessed by AQ score, tended to cooperate less with the ToMPlan agent than subjects with low levels of autistic traits. Conversely, these high AQ subjects tended to cooperate more with the Random agent than low AQ subjects. These findings suggest that high AQ subjects had more difficulty understanding the behavior and intentions of the ToMPlan agent, and were unable to form a cooperative equilibrium, often opting for hare hunting. Additionally, high AQ subjects, approximating the systemizing tendencies seen in clinically diagnosed individuals with ASD, appeared to be more adept at understanding the pattern of behavior exhibited by the Random agent, allowing them to capitalize on stag hunting at a higher frequency than low AQ subjects.

Chapter 4 discussed the pilot data for a proposed fMRI study that investigates the correlation between levels of autistic traits in general population individuals and differential neural expression while attempting social decision-making in a game theoretic environment. From the limited subject pool, trends were identified in several brain areas implicated in the ToM network. Subjects exhibited higher activation assessed by BOLD response in the rTPJ, right fusiform gyrus, and right medial temporal gyrus, when playing with an adaptive agent that simulated theory of mind. Furthermore, subjects with high levels of autistic traits showed less activation in rTPJ and rMT, brain areas that have been shown to process theory of mind assessments, and more activation in mPFC, an area related to higher level cognition and systematic decision-making, when compared to subjects with low levels of autistic traits. These trends indicate that subjects with high levels of autistic traits tend to recruit ToM areas less than subjects with less autistic traits, potentially increasing the recruitment of mPFC to compensate.

Taken together, this body of work contributes to our understanding of the differing abilities of individuals in the general population to respond in situations of varying levels of uncertainty. Here, we introduce two paradigms to investigate subject response to uncertainty in the presence of varying ability to form social contracts. A recurring theme linking the above studies was the success of the adaptive agent to engage complex and varied behavior from subjects. Moreover, the correlation of AQ score to cooperative strategizing that differs depending on the strategy used by the agent further suggests that using adaptive agents could have important implications for the study of ASD and other neurological disorders. These results warrant further investigation into their origins and the scope of their effects on behavior.

Additionally, extending this research from the general population as correlated by AQ score to clinical ASD populations could prove valuable. While it is generally accepted that autism is caused by structural or functional abnormalities in the brain, there is currently no known explicit point of origin for this anomalous development. As a direct result of the proposed research, deciphering exactly what types of ToM tasks people with high levels of autistic traits have trouble with (i.e., explicit vs. implicit), as well as the neural expression behind that deficit (i.e., hyper- vs. hypo-activation) will take the field closer to understanding the differences in underlying brain activity within areas recruited in the mentalizing network in people with ASD in order to develop targeted treatment methods and informed theories of autistic development at a population level.

Due to varying levels of ability in verbal comprehension and higher executive functioning among people with ASD, it is possible that the commonly used narrative-driven ToM tasks may be biased or gamed (Matthews et al., 2012) by cognitively able individuals who can compensate for ToM deficits. The proposed research utilizes the dynamic, non-verbal Stag Hunt task, a more neutral, implicit venue for investigating atypical ToM activation, less likely to be confounded by individual differences in compensatory strategies. This ensures the data collected will be more accurate and meaningful when applied to the context of differences between ASD and NT subjects.
The incorporation of an adaptive ToM agent will provide insight into differences in activation of ROIs that result from using an intelligent, non-human player that has of yet remained unexplored. As individuals with autism have been shown to respond well to social robotics (Scassellati et al., 2012), we believe that using adaptive agents in the proposed research is a natural addition to ASD research as intelligent, versatile interaction partners lacking common stressors present in human social partners, capable of probing complex behavior, social interaction, and ToM response.

Utilizing a dynamic game theoretic environment with fMRI subjects will provide valuable insight into what cues are important in determining whether another entity is worthy of ToM while recording in real-time the neural response to those entities. This research will provide a multimodal foundation for establishing a functional network of ToM that can be used to develop more intelligent technology, improve upon the understanding of ToM in the brain, and work towards understanding and treating disorders that affect ToM performance.

REFERENCES

Abu-Akel, A., & Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia*, *49*(11), 2971–2984. https://doi.org/10.1016/j.neuropsychologia.2011.07.012

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). Washington, D.C: American Psychiatric Association.

Anderhub, V., Engelmann, D., & Güth, W. (2002). An experimental study of the repeated trust game with incomplete information. *Journal of Economic Behavior & Organization*, *48*(2), 197–216.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, *16*(10), 1773–1784.

Asher, D. E., Barton, B., Brewer, A. A., & Krichmar, J. L. (2012). Reciprocity and retaliation in social games with adaptive agents. *IEEE Transactions on Autonomous Mental Development*, *4*(3), 226–238.

Asher, D. E., Zaldivar, A., & Krichmar, J. L. (2010). Effect of Neuromodulation on Performance in Game Playing: A Modeling Study. In *IEEE 9th International Conference on Development and Learning (ICDL)* (pp. 155–160).

Asher, D. E., Zhang, S., Zaldivar, A., Lee, M. D., & Krichmar, J. L. (2012). Modeling individual differences in socioeconomic game playing. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Sapporo, Japan.

Autism Spectrum Disorder. (2013). American Psychiatric Association. Retrieved from http://www.dsm5.org/Documents/Autism%20Spectrum%20Disorder%20Fact%20Sheet.pdf

Baldassarre, G. (2003). Forward and bidirectional planning based on reinforcement learning and neural networks in a simulated robot. In *Anticipatory Behavior in Adaptive Learning Systems* (pp. 179–200). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-45002-3_11

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 361–374. https://doi.org/10.1098/rstb.2002.1206

Baron-Cohen, Simon. (2009). Autism: The Empathizing-Systemizing (E-S) Theory. *Annals of the New York Academy of Sciences*, *1156*(1), 68–80. https://doi.org/10.1111/j.1749-6632.2009.04467.x

Baron-Cohen, Simon, Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37–46.

Baron-Cohen, Simon, Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience*, *11*(6), 1891–1898.

Baron-Cohen, Simon, Wheelwright, 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.

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, *7*(4), 404–410. https://doi.org/10.1038/nn1209

Beauchamp, M. S. (2015). The social mysteries of the superior temporal sulcus. *Trends in Cognitive Sciences*, *19*(9), 489–490. https://doi.org/10.1016/j.tics.2015.07.002

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with Autism Spectrum Disorders. *Autism*, *16*(4), 357–366.

Benoit, R. G., Szpunar, K. K., & Schacter, D. L. (2014). Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1419274111

Bevilacqua, L., & Goldman, D. (2011). Genetics of emotion. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.07.009

Bishara, A. J., Kruschke, J. K., Stout, J. C., Bechara, A., McCabe, D. P., & Busemeyer, J. R. (2010). Sequential learning models for the Wisconsin card sort task: Assessing processes in substance dependent individuals. *Journal of Mathematical Psychology*, *54*(1), 5–13. https://doi.org/10.1016/j.jmp.2008.10.002

Boraston, Z., & Blakemore, S.-J. (2007). The application of eye-tracking technology in the study of autism. *The Journal of Physiology*, *581*(3), 893–898. https://doi.org/10.1113/jphysiol.2007.133587

Bosse, T., Memon, Z. A., & Treur, J. (2007). A two-level BDI-agent model for theory of mind and its use in social manipulation. In *Proceedings of the AISB 2007 Workshop on Mindful Environments* (pp. 335–342). Retrieved from http://neuro.bstu.by/ai/To-dom/My_research/Papers-3/Intention/BDI-model/AISB07-mindfulBDI-final.pdf

Boucher, J. (2012). Putting theory of mind in its place: psychological explanations of the socio-emotional-communicative impairments in autistic spectrum disorder. *Autism*, *16*(3), 226–246. https://doi.org/10.1177/1362361311430403

Boureau, Y. L., & Dayan, P. (2010). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, *36*(1), 74–97.

Calderon, J. P. (2006). How Learning Affects the Evolution of Strong Reciprocity. *Adaptive Behavior*, *14*(3), 211–221. https://doi.org/10.1177/105971230601400306

Camille, N. (2004). The Involvement of the Orbitofrontal Cortex in the Experience of Regret. *Science*, *304*(5674), 1167–1170. https://doi.org/10.1126/science.1094550

Carretié, L., Mercado, F., Tapia, M., & Hinojosa, J. A. (2001). Emotion, attention, and the "negativity bias", studied through event-related potentials. *International Journal of Psychophysiology*, *41*(1), 75–85. https://doi.org/10.1016/S0167-8760(00)00195-1

Caspi, A. (2003). Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene. *Science*, *301*(5631), 386–389. https://doi.org/10.1126/science.1083968

Caspi, A., Hariri, A. R., Holmes, A., Uher, R., & Moffitt, T. E. (2010). Genetic Sensitivity to the Environment: The Case of the Serotonin Transporter Gene and Its Implications for Studying Complex Diseases and Traits. *American Journal of Psychiatry*, *167*(5), 509–527. https://doi.org/10.1176/appi.ajp.2010.09101452

Castelli, F. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839–1849. https://doi.org/10.1093/brain/awf189

Chaminade, T., Rosset, D., Da Fonseca, D., Nazarian, B., Lutcher, E., Cheng, G., & Deruelle, C. (2012). How do we think machines think? An fMRI study of alleged competition with an artificial intelligence. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00103

Chevallier, C., Parish-Morris, J., Tonge, N., Le, L., Miller, J., & Schultz, R. T. (2014). Susceptibility to the audience effect explains performance gap between children with and without autism in a theory of mind task. *Journal of Experimental Psychology: General*, *143*(3), 972–979. https://doi.org/10.1037/a0035483

Ciaramelli, E., Bernardi, F., & Moscovitch, M. (2013). Individualized Theory of Mind (iToM): When Memory Modulates Empathy. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00004

Congdon, E., Lesch, K. P., & Canli, T. (2008). Analysis of DRD4 and DAT polymorphisms and behavioral inhibition in healthy adults: Implications for impulsivity. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *147B*(1), 27–32. https://doi.org/10.1002/ajmg.b.30557

Cools, R., Nakamura, K., & Daw, N. D. (2010). Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology*, *36*(1), 98–113.

Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A., & Dolan, R. J. (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nature Neuroscience*, *8*(9), 1255–1262. https://doi.org/10.1038/nn1514

Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*, *121*(3), 463.

Craig, A. B., Asher, D. E., Oros, N., Brewer, A. A., & Krichmar, J. L. (2013). Social contracts and human-computer interaction with simulated adapting agents. *Adaptive Behavior*.

Craig, Alexis B., Phillips, M. E., Zaldivar, A., Bhattacharyya, R., & Krichmar, J. L. (2016). Investigation of Biases and Compensatory Strategies Using a Probabilistic Variant of the Wisconsin Card Sorting Test. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00017

Dayan, P. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Cambridge, Mass.: Massachusetts Institute of Technology Press.

Deen, B., & Saxe, R. R. (2012). Neural correlates of social perception: The posterior superior temporal sulcus is modulated by action rationality, but not animacy. In *Proceedings of the 33rd Annual Cognitive Science Society Conference* (pp. 276–81). Retrieved from http://mindmodeling.org/cogsci2012/papers/0060/paper0060.pdf

Dehaene, S., & Changeux, J.-P. (1991). The Wisconsin Card Sorting Test: theoretical analysis and modeling in a neuronal network. *Cerebral Cortex*, *1*(1), 62–79.

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic Genes Predict Individual Differences in Susceptibility to Confirmation Bias. *Journal of Neuroscience*, *31*(16), 6188–6198. https://doi.org/10.1523/JNEUROSCI.6486-10.2011

Eigsti, I.-M., & Shapiro, T. (2003). A systems neuroscience approach to autism: Biological, cognitive, and clinical perspectives. *Mental Retardation and Developmental Disabilities Research Reviews*, *9*(3), 206–216. https://doi.org/10.1002/mrdd.10081

El Kaliouby, R., & Robinson, P. (2005). Generalization of a vision-based computational model of mind-reading. In *Affective computing and intelligent interaction* (pp. 582–589). Springer. Retrieved from http://link.springer.com/chapter/10.1007/11573548_75

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, *101*(3), 519–527. https://doi.org/10.1037/0033-295X.101.3.519

Feil-Seifer, D., & Matarić, M. J. (2009). Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. In *Experimental robotics* (pp. 201–210). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-00196-3_24

Fine, C. (2001). Dissociation between `theory of mind' and executive functions in a patient with early left amygdala damage. *Brain*, *124*(2), 287–298. https://doi.org/10.1093/brain/124.2.287

Fleck, M. S., Samei, E., & Mitroff, S. R. (2010). Generalized "satisfaction of search": Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied*, *16*(1), 60.

Focquaert, F., & Vanneste, S. (2015). Autism spectrum traits in normal individuals: a preliminary VBM analysis. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00264

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. *Journal of Neuroscience*, *35*(2), 485–494. https://doi.org/10.1523/JNEUROSCI.2036-14.2015

Friedlander, D., & Franklin, S. (2008). LIDA and a Theory of Mind. In *Artificial General Intelligence, 2008: Proceedings of the First AGI Conference* (Vol. 171, p. 137). IOS Press.

Frith, C. D., & Frith, U. (2008). Implicit and Explicit Processes in Social Cognition. *Neuron*, *60*(3), 503–510. https://doi.org/10.1016/j.neuron.2008.10.032

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, *109*(3), 416–422. https://doi.org/10.1016/j.cognition.2008.09.007

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind." *Trends in Cognitive Sciences*, *7*(2), 77–83.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of "theory of mind"in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases." *European Review of Social Psychology*, *2*(1), 83–115.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, *103*(4), 650.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Supplement_3), 15647–15654. https://doi.org/10.1073/pnas.1014269108

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*(11), 1803–1814.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404.

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*, *105*(18), 6741–6746.

Hari, R., & Kujala, M. V. (2009). Brain Basis of Human Social Interaction: From Concepts to Brain Imaging. *Physiological Reviews*, *89*(2), 453–479. https://doi.org/10.1152/physrev.00041.2007

Hariri, A. R. (2002). Serotonin Transporter Genetic Variation and the Response of the Human Amygdala. *Science*, *297*(5580), 400–403. https://doi.org/10.1126/science.1071829

Haselton, M., Nettle, D., & Andrews, P. (2005). The evolution of cognitive bias. In *The handbook of evolutionary psychology* (pp. 724–746).

Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Sciences*, *18*(4), 167–170. https://doi.org/10.1016/j.tics.2014.01.005

Hilbert, M. (2012). Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin*, *138*(2), 211–237. https://doi.org/10.1037/a0025940

Hoehn, B., Southey, F., Holte, R. C., & Bulitko, V. (2005). Effective short-term opponent exploitation in simplified poker. In *AAAI* (Vol. 5, pp. 783–788). Retrieved from http://www.aaai.org/Papers/AAAI/2005/AAAI05-123.pdf

Homberg, J. R., & Lesch, K.-P. (2011). Looking on the Bright Side of Serotonin Transporter Gene Variation. *Biological Psychiatry*, *69*(6), 513–519. https://doi.org/10.1016/j.biopsych.2010.09.024

Howes, A., Duggan, G. B., Kalidindi, K., Tseng, Y.-C., & Lewis, R. L. (2015). Predicting Short-Term Remembering as Boundedly Optimal Strategy Choice. Retrieved from http://www-personal.umich.edu/~rickl/pubs/howes-duggan-kalidindi-tseng-2015-cogsci.pdf

Hyde, L. W., Bogdan, R., & Hariri, A. R. (2011). Understanding risk for psychopathology through imaging gene–environment interactions. *Trends in Cognitive Sciences*, *15*(9), 417–427. https://doi.org/10.1016/j.tics.2011.07.001

Imhof, L. A., Fudenberg, D., & Nowak, M. A. (2007). Tit-for-tat or win-stay, lose-shift? *Journal of Theoretical Biology*, *247*(3), 574–580.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*(4), 887–900. https://doi.org/10.1037/0022-3514.75.4.887

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Social Neuroscience*, *4*(2), 135–152. https://doi.org/10.1080/17470910802198510

Kana, R. K., Keller, T. A., Minshew, N. J., & Just, M. A. (2007). Inhibitory Control in High-Functioning Autism: Decreased Activation and Underconnectivity in Inhibition Networks. *Biological Psychiatry*, *62*(3), 198–206. https://doi.org/10.1016/j.biopsych.2006.08.004

Khamassi, M., Lachèze, L., Girard, B., Berthoz, A., & Guillot, A. (2005). Actor–Critic Models of Reinforcement Learning in the Basal Ganglia: From Natural to Artificial Rats. *Adaptive Behavior*, *13*(2), 131–148. https://doi.org/10.1177/105971230501300205

Kidd, C. D., & Breazeal, C. (2004). Effect of a robot on user perceptions. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on* (Vol. 4, pp. 3559–3564). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1389967

Kiesler, S., Sproull, L., & Waters, K. (1996). A prisoner's dilemma experiment on cooperation with people and human-like computers. *Journal of Personality and Social Psychology*, *70*(1), 47.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation*, *32*, 385–418.

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, *110*(14), 5648–5653. https://doi.org/10.1073/pnas.1207992110

Krichmar, J. L. (2008). The Neuromodulatory System: A Framework for Survival and Adaptive Behavior in a Challenging World. *Adaptive Behavior*, *16*(6), 385–399. https://doi.org/10.1177/1059712308095775

Lee, D. (2008). Game theory and neural basis of social decision making. *Nature Neuroscience*, *11*(4), 404–409. https://doi.org/10.1038/nn2065

Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, *15*(1), 1–15. https://doi.org/10.3758/PBR.15.1.1

Lee, Michael D., Zhang, S., Munro, M., & Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research*, *12*(2), 164–174. https://doi.org/10.1016/j.cogsys.2010.07.007

Lesch, K.-P., Bengel, D., Heils, A., Sabol, S. Z., Greenberg, B. D., Petri, S., … Murphy, D. L. (1996). Association of Anxiety-Related Traits with a Polymorphism in the Serotonin Transporter Gene Regulatory Region. *Science*, *274*(5292), 1527–1531. https://doi.org/10.1126/science.274.5292.1527

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in “theory of mind.” *Trends in Cognitive Sciences*, *8*(12), 528–533. https://doi.org/10.1016/j.tics.2004.10.001

Li, C., Lowe, R., & Ziemke, T. (2013). Humanoids Learning to Walk: A Natural CPG-Actor-Critic Architecture. *Frontiers in Neurorobotics*, *7*. https://doi.org/10.3389/fnbot.2013.00005

Lie, C.-H., Specht, K., Marshall, J. C., & Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *NeuroImage*, *30*(3), 1038–1049. https://doi.org/10.1016/j.neuroimage.2005.10.031

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Sadek, S. A., Pasco, G., Wheelwright, S. J., … Baron-Cohen, S. (2010). Atypical neural self-representation in autism. *Brain*, *133*(2), 611–624. https://doi.org/10.1093/brain/awp306

Lombardo, Michael V., Chakrabarti, B., Bullmore, E. T., & Baron-Cohen, S. (2011). Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *NeuroImage*, *56*(3), 1832–1838. https://doi.org/10.1016/j.neuroimage.2011.02.067

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Loth, E., Carvalho, F., & Schumann, G. (2011). The contribution of imaging genetics to the development of predictive markers for addictions. *Trends in Cognitive Sciences*, *15*(9), 436–446. https://doi.org/10.1016/j.tics.2011.07.008

Mars, R. B., Neubert, F.-X., Noonan, M. P., Sallet, J., Toni, I., & Rushworth, M. F. S. (2012). On the relationship between the “default mode network” and the “social brain.” *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00189

Martin, A., & Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cognitive Neuropsychology*, *20*(3-6), 575–587.

Mason, R. A., Williams, D. L., Kana, R. K., Minshew, N., & Just, M. A. (2008). Theory of Mind disruption and recruitment of the right hemisphere during narrative comprehension in autism. *Neuropsychologia*, *46*(1), 269–280. https://doi.org/10.1016/j.neuropsychologia.2007.07.018

Matthews, N. L., Goldberg, W. A., Lukowski, A. F., Osann, K., Abdullah, M. M., Ly, A. R., … Anne Spence, M. (2012). Does theory of mind performance differ in children with early-onset and regressive autism?: Theory of mind in early-onset and regressive autism. *Developmental Science*, *15*(1), 25–34. https://doi.org/10.1111/j.1467-7687.2011.01094.x

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge; New York: Cambridge University Press.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832–11835.

Merrick, K. E. (2010). A Comparative Study of Value Systems for Self-Motivated Exploration and Learning by Robots. *IEEE Transactions on Autonomous Mental Development*, *2*(2), 119–131. https://doi.org/10.1109/TAMD.2010.2051435

Merrick, K., & Maher, M. L. (2009). Motivated Learning from Interesting Events: Adaptive, Multitask Learning Agents for Complex Environments. *Adaptive Behavior*, *17*(1), 7–27. https://doi.org/10.1177/1059712308100236

Milner, B. (1963). Effects of Different Brain Lesions on Card Sorting: The Role of the Frontal Lobes. *Archives of Neurology*, *9*(1), 90. https://doi.org/10.1001/archneur.1963.00460070100010

Moll, J., de Oliveira-Souza, R., Bramati, I. E., & Grafman, J. (2002). Functional Networks in Emotional Moral and Nonmoral Social Judgments. *NeuroImage*, *16*(3), 696–703. https://doi.org/10.1006/nimg.2002.1118

Munafò, M. R., Yalcin, B., Willis-Owen, S. A., & Flint, J. (2008). Association of the Dopamine D4 Receptor (DRD4) Gene and Approach-Related Personality Traits: Meta-Analysis and New Data. *Biological Psychiatry*, *63*(2), 197–206. https://doi.org/10.1016/j.biopsych.2007.04.006

Nakamura, Y., Mori, T., Sato, M., & Ishii, S. (2007). Reinforcement learning for a biped robot based on a CPG-actor-critic method. *Neural Networks*, *20*(6), 723–735. https://doi.org/10.1016/j.neunet.2007.01.002

Navarro, D. J., & Newell, B. (2014). Information versus reward in a changing world. In *36th Annual Meeting of the Cognitive Science Society (CogSci 2014)(23 Jul 2014-26 Jul 2014: Quebec City, Canada)*. Retrieved from https://digital.library.adelaide.edu.au/dspace/handle/2440/91081

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness Versus Reason in the Ultimatum Game. *Science*, *289*(5485), 1773–1775. https://doi.org/10.1126/science.289.5485.1773

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*(6432), 56–58. https://doi.org/10.1038/364056a0

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*, *71*(3), 437–451. https://doi.org/10.1016/j.bandc.2009.03.005

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, *130*(7), 1718–1731. https://doi.org/10.1093/brain/awm052

Phillips, M. E., Chelian, S. E., Pirolli, P., & Bhattacharyya, R. (2014). Forensic foraging of change detection in opponent strategies with a neural model of the interactions between temporal and prefrontal cortex. *Biologically Inspired Cognitive Architectures*, *10*, 17–23. https://doi.org/10.1016/j.bica.2014.11.003

Pynadath, D. V., & Marsella, S. C. (2005). PsychSim: Modeling theory of mind with decision-theoretic agents. In *IJCAI* (Vol. 5, pp. 1181–1186). Retrieved from http://ijcai.org/Past%20Proceedings/IJCAI-05/PDF/1559.pdf

Rand, D. G., Tarnita, C. E., Ohtsuki, H., & Nowak, M. A. (2013). From the Cover: Evolution of fairness in the one-shot anonymous Ultimatum Game. *Proceedings of the National Academy of Sciences*, *110*(7), 2581–2586. https://doi.org/10.1073/pnas.1214167110

Rilling, J. K., & Sanfey, A. G. (2011). The Neuroscience of Social Decision-Making. *Annual Review of Psychology*, *62*(1), 23–48. https://doi.org/10.1146/annurev.psych.121208.131647

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*(4), 1694–1703. https://doi.org/10.1016/j.neuroimage.2004.04.015

Rissanen, J. (1983). A Universal Prior for Integers and Estimation by Minimum Description Length. *The Annals of Statistics*, *11*(2), 416–431. https://doi.org/10.1214/aos/1176346150

Robinson, A. L., Heaton, R. K., Lehman, R. A., & Stilson, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of Consulting and Clinical Psychology*, *48*(5), 605.

Rolison, M. J., Naples, A. J., & McPartland, J. C. (2015). Interactive Social Neuroscience to Study Autism Spectrum Disorder. *The Yale Journal of Biology and Medicine*, *88*(1), 17.

Rougier, N. P., & O'Reilly, R. C. (2002). Learning representations in a gated prefrontal cortex model of dynamic task switching. *Cognitive Science*, *26*(4), 503–520.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320.

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism*, *6*(1), 2.

Sabbagh, M. A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism. *Brain and Cognition*, *55*(1), 209–219. https://doi.org/10.1016/j.bandc.2003.04.002

Sanfey, A. G. (2003). The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science*, *300*(5626), 1755–1758. https://doi.org/10.1126/science.1082976

Sang, K., Todd, P. M., & Goldstone, R. L. (2011). Learning near-optimal search in a minimal explore/exploit task. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2800–2805). Retrieved from http://mindmodeling.org/cogsci2011/papers/0642/paper0642.pdf

Saxe, R, & Kanwisher, N. (2003). People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R., & Powell, L. J. (2006). It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind. *Psychological Science*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Saxe, Rebecca. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*(2), 235–239. https://doi.org/10.1016/j.conb.2006.03.001

Scassellati, B., Henny Admoni, & Matarić, M. (2012). Robots for Use in Autism Research. *Annual Review of Biomedical Engineering*, *14*(1), 275–294. https://doi.org/10.1146/annurev-bioeng-071811-150036

Schembri, M., Mirolli, M., & Baldassarre, G. (2007). Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot. In *IEEE 6th International Conference on Development and Learning, 2007. ICDL 2007* (pp. 282 –287). https://doi.org/10.1109/DEVLRN.2007.4354052

Scholz, J. T., & Whiteman, M. A. (2010). Social Capital in Coordination Experiments: Risk, Trust and Position. Retrieved from http://opensiuc.lib.siu.edu/pn_wp/50/

Schultz, M. T., Mitchell, K. N., Harper, B. K., & Bridges, T. S. (2010). *Decision making under uncertainty*. DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA534878

Schuwerk, T., Vuori, M., & Sodian, B. (2014). Implicit and explicit Theory of Mind reasoning in autism spectrum disorders: The impact of experience. *Autism*. https://doi.org/10.1177/1362361314526004

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion*, *14*(4), 433–440. https://doi.org/10.1080/026999300402745

Senju, A. (2012). Spontaneous Theory of Mind and Its Absence in Autism Spectrum Disorders. *The Neuroscientist*, *18*(2), 108–113. https://doi.org/10.1177/1073858410397208

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind Eyes: An Absence of Spontaneous Theory of Mind in Asperger Syndrome. *Science*, *325*(5942), 883–885. https://doi.org/10.1126/science.1176170

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, *15*(3), 233–250. https://doi.org/10.1002/bdm.413

Si, M., Marsella, S. C., & Pynadath, D. V. (2010). Modeling appraisal in theory of mind reasoning. *Autonomous Agents and Multi-Agent Systems*, *20*(1), 14–31. https://doi.org/10.1007/s10458-009-9093-x

Simon, H. (1972). Theories of bounded rationality. *Decision and Organization*, *1*(1), 161–176.

Simons, D. J. (2010). Monkeying around with the gorillas in our midst: familiarity with an inattentional-blindness task does not improve the detection of unexpected events. *i-Perception*, *1*(1), 3–6. https://doi.org/10.1068/i0386

Skyrms, B. (2001). The Stag Hunt. Presented at the Presidential Address Pacific Division of the American Philosophical Association.

Skyrms, B., & Pemantle, R. (2000). A dynamic model of social network formation. *Proceedings of the National Academy of Sciences*, *97*(16), 9340–9346.

Skyrms, Brian. (2004). *The stag hunt and the evolution of social structure*. Cambridge, UK; New York: Cambridge University Press.

Spek, A. A., Scholte, E. M., & Berckelaer-Onnes, I. A. (2010). Theory of Mind in Adults with HFA and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, *40*(3), 280–289. https://doi.org/10.1007/s10803-009-0860-y

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (2013). Theory of Mind. *Autism: The Science of Mental Health*, 226.

Sutton, R. S., Barto, A. G., & Williams, R. J. (1992). Reinforcement learning is direct adaptive optimal control. *IEEE Control Systems*, *12*(2), 19 –22. https://doi.org/10.1109/37.126844

Szolnoki, A., & Perc, M. (2009). Resolving social dilemmas on evolving random networks. *EPL (Europhysics Letters)*, *86*, 30007.

Szolnoki, Attila, Xie, N.-G., Ye, Y., & Perc, M. (2013). Evolution of emotions on networks leads to the evolution of cooperation in social dilemmas. *Physical Review E*, *87*(4). https://doi.org/10.1103/PhysRevE.87.042805

Thomaz, A L, Berlin, M., & Breazeal, C. (2005). Robot science meets social science: An embodied computational model of social referencing. In *Workshop toward social mechanisms of android science* (pp. 7–17).

Thomaz, Andrea L., & Breazeal, C. (2008). Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence*, *172*(6-7), 716–737. https://doi.org/10.1016/j.artint.2007.09.009

Thomaz, Andrea Lockerd, Berlin, M., & Breazeal, C. (2005). An embodied computational model of social referencing. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on* (pp. 591–598). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1513844

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1130.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Valluri, A. (2006). Learning and Cooperation in Sequential Games. *Adaptive Behavior*, *14*(3), 195–209. https://doi.org/10.1177/105971230601400304

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547

Von dem Hagen, E. A. H., Nummenmaa, L., Yu, R., Engell, A. D., Ewbank, M. P., & Calder, A. J. (2011). Autism Spectrum Traits in the Typical Population Predict Structure and Function in the Posterior Superior Temporal Sulcus. *Cerebral Cortex*, *21*(3), 493–500. https://doi.org/10.1093/cercor/bhq062

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101–118.

Wang, H., & Doering, L. C. (2015). Autism spectrum disorders: emerging mechanisms and mechanism-based treatment. *Frontiers in Cellular Neuroscience*, *9*. https://doi.org/10.3389/fncel.2015.00183

Wei, Z., Zhao, Z., & Zheng, Y. (2013). Neural mechanisms underlying social conformity in an ultimatum game. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00896

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., … Wakabayashi, A. (2006). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). *Brain Research*, *1079*(1), 47–56. https://doi.org/10.1016/j.brainres.2006.01.012

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development*, *80*(4), 1097–1117.

White, S. J., Frith, U., Rellecke, J., Al-Noor, Z., & Gilbert, S. J. (2014). Autistic adolescents show atypical activation of the brain′s mentalizing system even without a prior history of mentalizing problems. *Neuropsychologia*, *56*, 17–25. https://doi.org/10.1016/j.neuropsychologia.2013.12.013

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081. https://doi.org/10.1037/a0038199

Wilson, R. C., & Niv, Y. (2012). Inferring Relevance in a Changing World. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00189

Wood, R. M., Rilling, J. K., Sanfey, A. G., Bhagwagar, Z., & Rogers, R. D. (2006). Effects of Tryptophan Depletion on the Performance of an Iterated Prisoner's Dilemma Game in Healthy Adults. *Neuropsychopharmacology*, *31*(5), 1075–1084. https://doi.org/10.1038/sj.npp.1300932

Worthy, D. A., Maddox, W. T., & Markman, A. B. (2007). Regulatory fit effects in a choice task. *Psychonomic Bulletin & Review*, *14*(6), 1125–1132.

Wout, M., Kahn, R. S., Sanfey, A. G., & Aleman, A. (2006). Affective state and decision-making in the Ultimatum Game. *Experimental Brain Research*, *169*(4), 564–568. https://doi.org/10.1007/s00221-006-0346-5

Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability Matching as a Computational Strategy Used in Perception. *PLoS Computational Biology*, *6*(8), e1000871. https://doi.org/10.1371/journal.pcbi.1000871

Yoshida, W., Dziobek, I., Kliemann, D., Heekeren, H. R., Friston, K. J., & Dolan, R. J. (2010). Cooperation and Heterogeneity of the Autistic Mind. *Journal of Neuroscience*, *30*(26), 8815–8818. https://doi.org/10.1523/JNEUROSCI.0400-10.2010

Yoshida, W., Seymour, B., Friston, K. J., & Dolan, R. J. (2010). Neural Mechanisms of Belief Inference during Cooperative Games. *Journal of Neuroscience*, *30*(32), 10744–10751. https://doi.org/10.1523/JNEUROSCI.5895-09.2010

Yoshida, Wako, Dolan, R. J., & Friston, K. J. (2008). Game Theory of Mind. *PLoS Computational Biology*, *4*(12), e1000254. https://doi.org/10.1371/journal.pcbi.1000254

Young, K. A., Holcomb, L. A., Bonkale, W. L., Hicks, P. B., Yazdani, U., & German, D. C. (2007). 5HTTLPR Polymorphism and Enlargement of the Pulvinar: Unlocking the Backdoor to the Limbic System. *Biological Psychiatry*, *61*(6), 813–818. https://doi.org/10.1016/j.biopsych.2006.08.047

Yu, A. J., & Dayan, P. (2005). Uncertainty, Neuromodulation, and Attention. *Neuron*, *46*(4), 681–692. https://doi.org/10.1016/j.neuron.2005.04.026

Zagare, F. C. (1984). *Game theory : concepts and applications*. Beverly Hills: Sage Publications.

Zaldivar, A., Asher, D., & Krichmar, J. (2010). Simulation of how neuromodulation influences cooperative behavior. *From Animals to Animats 11*, 649–660.

Zalla, T., Miele, D., Leboyer, M., & Metcalfe, J. (2015). Metacognition of agency and theory of mind in adults with high functioning autism. *Consciousness and Cognition*, *31*, 126–138. https://doi.org/10.1016/j.concog.2014.11.001

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Score – Blockwise Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by KS test statistic.

Block	1	2	3	4	5	6	7	8	9	10	11
1	0.000	0.317	0.667	0.733	0.700	0.500	0.250	0.717	0.733	0.200	0.367
2		0.000	0.533	0.750	0.533	0.350	0.550	0.583	0.733	0.483	0.167
3			0.000	0.400	0.183	0.417	0.767	0.183	0.250	0.750	0.700
4				0.000	0.533	0.667	0.850	0.533	0.167	0.867	0.850
5					0.000	0.367	0.800	0.133	0.433	0.767	0.683
6						0.000	0.717	0.417	0.617	0.683	0.333
7							0.000	0.800	0.850	0.117	0.617
8								0.000	0.400	0.783	0.733
9									0.000	0.867	0.833
10										0.000	0.533
11											0.000

Supplementary Table 2. Score – Blockwise Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by p-value.

Block	1	2	3	4	5	6	7	8	9	10	11
1	1.000	0.004	<0.001	<0.001	<0.001	<0.001	0.039	<0.001	<0.001	0.160	<0.001
2		1.000	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.345
3			1.000	<0.001	0.239	<0.001	<0.001	0.239	0.039	<0.001	<0.001
4				1.000	<0.001	<0.001	<0.001	<0.001	0.345	<0.001	<0.001
5					1.000	<0.001	<0.001	0.629	<0.001	<0.001	<0.001
6						1.000	<0.001	<0.001	<0.001	<0.001	0.002
7							1.000	<0.001	<0.001	0.784	<0.001
8								1.000	<0.001	<0.001	<0.001
9									1.000	<0.001	<0.001
10										1.000	<0.001
11											1.000

Supplementary Table 3. Score – Blockwise mean and standard deviation (in points).

Block	Mean	SD
1	34.667	41.213
2	28.433	28.727
3	-2.867	19.472
4	-19.900	16.099
5	-1.300	21.201
6	14.967	22.439
7	52.200	30.480
8	-0.500	16.367
9	-16.600	17.335
10	51.600	27.651
11	31.233	21.691

Supplementary Table 4. Score – Uncertainty level Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by KS test statistic.

Uncertainty	No	Low	Mod	High
No	0.000	0.494	0.733	0.817
Low		0.000	0.522	0.733
Mod			0.000	0.400
High				0.000

Supplementary Table 5. Score – Uncertainty level Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by p-value.

Uncertainty	No	Low	Mod	High
No	1	<0.001	<0.001	<0.001
Low		1	<0.001	<0.001
Mod			1	<0.001
High				1

Supplementary Table 6. Score – Uncertainty level mean and standard deviation (in points).

Uncertainty	Mean	SD
No	46.156	34.417
Low	24.878	25.372
Mod	-1.556	19.037
High	-18.250	16.740

Supplementary Table 7. Observe use (all trials) – Uncertainty level Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty levels as reported by KS test statistic.

Uncertainty	No	Low	Mod	High
No	0.000	0.086	0.197	0.300
Low		0.000	0.117	0.214
Mod			0.000	0.125
High				0.000

Supplementary Table 8. Observe use (all trials) – Uncertainty level Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty levels as reported by p-value.

Uncertainty	No	Low	Mod	High
No	1.000	0.640	0.006	<0.001
Low		1.000	0.161	0.002
Mod			1.000	0.196
High				1.000

Supplementary Table 9. Observe use (all trials) – Uncertainty level mean and standard deviation (in # observes).

Uncertainty	Mean [#]	SD
No	9.400	13.710
Low	9.706	12.855
Mod	11.789	13.219
High	15.042	14.908

Supplementary Table 10. Observe use (by half block). The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty levels, mean and standard deviation of each half block (in proportion of trials using observe). The first block did not feature the option to observe.

Block	Mean [1st]	Mean [2nd]	SD [1st]	SD [2nd]	P-value	KS-statistic
1	N/A	N/A	N/A	N/A	N/A	N/A
2	0.092	0.114	0.162	0.199	0.911	0.100
3	0.154	0.189	0.216	0.243	0.784	0.117
4	0.243	0.287	0.284	0.298	0.629	0.133
5	0.254	0.272	0.293	0.298	0.981	0.083
6	0.258	0.219	0.290	0.279	0.911	0.100
7	0.157	0.181	0.261	0.279	0.981	0.083
8	0.251	0.295	0.288	0.302	0.911	0.100
9	0.329	0.345	0.334	0.334	0.911	0.100
10	0.211	0.203	0.293	0.297	0.999	0.067
11	0.244	0.238	0.306	0.307	0.999	0.067

Supplementary Table 11. Run of Observes – Blockwise Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by KS test statistic. The first block did not feature the option to observe.

Block	1	2	3	4	5	6	7	8	9	10	11
1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2		0.000	0.063	0.308	0.348	0.407	0.351	0.392	0.332	0.295	0.360
3			0.000	0.301	0.346	0.348	0.291	0.366	0.307	0.297	0.335
4				0.000	0.127	0.140	0.084	0.164	0.104	0.055	0.135
5					0.000	0.119	0.164	0.101	0.133	0.128	0.154
6						0.000	0.138	0.109	0.130	0.140	0.188
7							0.000	0.173	0.131	0.121	0.159
8								0.000	0.102	0.113	0.146
9									0.000	0.107	0.126
10										0.000	0.154
11											0.000

Supplementary Table 12. Run of Observes – Blockwise Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by p-value. The first block did not feature the option to observe.

Block	1	2	3	4	5	6	7	8	9	10	11
1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2		1.000	1.000	0.029	0.015	0.002	0.022	0.003	0.018	0.058	0.011
3			1.000	0.024	0.011	0.010	0.071	0.005	0.026	0.043	0.015
4				1.000	0.853	0.751	0.999	0.549	0.953	1.000	0.794
5					1.000	0.929	0.690	0.982	0.832	0.884	0.708
6						1.000	0.864	0.961	0.849	0.805	0.443
7							1.000	0.612	0.888	0.946	0.731
8								1.000	0.973	0.950	0.756
9									1.000	0.964	0.873
10										1.000	0.708
11											1.000

Supplementary Table 13. Run of Observes – Blockwise mean and standard deviation (in average # of observes per run). The first block did not feature the option to observe.

Block	Mean [length]	SD
1	N/A	N/A
2	3.469	8.093
3	2.571	3.011
4	3.842	4.613
5	3.708	4.158
6	4.224	7.022
7	4.423	5.818

8	5.189	10.478
9	5.678	10.225
10	5.506	10.797
11	6.544	13.136

Supplementary Table 14. Run of Observes – Uncertainty level Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by KS test statistic. The first block did not feature the option to observe.

Uncertainty	No	Low	Mod	High
No	0.000	0.181	0.138	0.093
Low		0.000	0.202	0.187
Mod			0.000	0.127
High				0.000

Supplementary Table 15. Run of Observes – Uncertainty level Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by p-value. The first block did not feature the option to observe.

Uncertainty	No	Low	Mod	High
No	1.000	0.780	0.927	0.998
Low		1.000	0.652	0.681
Mod			1.000	0.934
High				1.000

Supplementary Table 16. Run of Observes – Uncertainty level mean and standard deviation (in average # of observes per run). The first block did not feature the option to observe.

Uncertainty	Mean	SD
No	4.673	6.959
Low	3.702	3.732
Mod	3.537	3.651
High	4.710	5.789

Supplementary Table 17. Win Stay – Blockwise Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by KS test statistic.

Block	1	2	3	4	5	6	7	8	9	10	11
1	0.000	0.359	0.224	0.427	0.333	0.233	0.300	0.317	0.283	0.291	0.371
2		0.000	0.220	0.475	0.428	0.197	0.081	0.127	0.262	0.109	0.097
3			0.000	0.339	0.229	0.097	0.232	0.147	0.107	0.279	0.256
4				0.000	0.127	0.295	0.506	0.403	0.265	0.520	0.538
5					0.000	0.267	0.444	0.353	0.233	0.487	0.464
6						0.000	0.239	0.162	0.083	0.268	0.260

7							0.000	0.155	0.276	0.068	0.071
8								0.000	0.190	0.185	0.162
9									0.000	0.308	0.287
10										0.000	0.098
11											0.000

Supplementary Table 18. Win Stay – Blockwise Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by p-value.

Block	1	2	3	4	5	6	7	8	9	10	11
1	1.000	<0.001	0.087	<0.001	0.002	0.064	0.007	0.004	0.014	0.011	<0.001
2		1.000	0.098	<0.001	<0.001	0.177	0.988	0.708	0.029	0.865	0.938
3			1.000	0.002	0.075	0.928	0.073	0.519	0.873	0.017	0.038
4				1.000	0.693	0.009	<0.001	<0.001	0.026	<0.001	<0.001
5					1.000	0.022	<0.001	<0.001	0.069	<0.001	<0.001
6						1.000	0.058	0.389	0.984	0.024	0.033
7							1.000	0.454	0.019	0.999	0.998
8								1.000	0.222	0.252	0.411
9									1.000	0.006	0.014
10										1.000	0.939
11											1.000

Supplementary Table 19. Win Stay – Blockwise mean and standard deviation (in proportion of stay trials after a winning trial).

Block	Mean	SD
1	0.846	0.254
2	0.890	0.211
3	0.862	0.216
4	0.732	0.257
5	0.745	0.281
6	0.832	0.243
7	0.896	0.218
8	0.903	0.184
9	0.833	0.238
10	0.920	0.171
11	0.918	0.201

Supplementary Table 20. Win Stay – Uncertainty level Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by KS test statistic.

Uncertainty	No	Low	Mod	High
No	0.000	0.131	0.228	0.361
Low		0.000	0.210	0.310
Mod			0.000	0.162
High				0.000

Supplementary Table 21. Win Stay – Uncertainty level Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by p-value.

Uncertainty	No	Low	Mod	High
No	1.000	0.089	<0.001	<0.001
Low		1.000	<0.001	<0.001
Mod			1.000	0.045
High				1.000

Supplementary Table 22. Win Stay – Uncertainty level mean and standard deviation (in proportion of stay trials after a winning trial).

Uncertainty	Mean	SD
No	0.887	0.219
Low	0.879	0.221
Mod	0.835	0.239
High	0.782	0.252

Supplementary Table 23. Lose Shift – Blockwise Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by KS test statistic.

Block	1	2	3	4	5	6	7	8	9	10	11
1	0.000	0.200	0.150	0.217	0.217	0.180	0.144	0.198	0.215	0.119	0.131
2		0.000	0.267	0.367	0.350	0.363	0.167	0.331	0.348	0.169	0.222
3			0.000	0.200	0.167	0.163	0.294	0.125	0.188	0.206	0.100
4				0.000	0.117	0.119	0.310	0.092	0.173	0.302	0.228
5					0.000	0.070	0.263	0.086	0.156	0.254	0.224
6						0.000	0.259	0.108	0.136	0.298	0.229
7							0.000	0.293	0.277	0.088	0.199
8								0.000	0.132	0.260	0.204
9									0.000	0.281	0.226
10										0.000	0.130
11											0.000

Supplementary Table 24. Lose Shift – Blockwise Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between blocks as reported by p-value.

Block	1	2	3	4	5	6	7	8	9	10	11

1	1.000	0.160	0.477	0.103	0.103	0.263	0.543	0.177	0.112	0.794	0.679
2		1.000	0.022	<0.001	<0.001	<0.001	0.351	0.002	0.001	0.363	0.104
3			1.000	0.160	0.345	0.374	0.009	0.720	0.218	0.159	0.926
4				1.000	0.784	0.770	0.005	0.956	0.309	0.009	0.089
5					1.000	0.998	0.028	0.975	0.432	0.042	0.099
6						1.000	0.032	0.862	0.618	0.010	0.089
7							1.000	0.011	0.018	0.976	0.193
8								1.000	0.659	0.037	0.171
9									1.000	0.018	0.097
10										1.000	0.725
11											1.000

Supplementary Table 25. Lose Shift – Blockwise mean and standard deviation (in proportion of shift trials after a losing trial).

Block	Mean	SD
1	0.748	0.174
2	0.801	0.168
3	0.719	0.166
4	0.694	0.126
5	0.680	0.177
6	0.695	0.157
7	0.763	0.206
8	0.695	0.159
9	0.659	0.193
10	0.745	0.208
11	0.722	0.198

Supplementary Table 26. Lose Shift – Uncertainty level Komolgorov Smirnov test statistic values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by KS test statistic.

Uncertainty	No	Low	Mod	High
No	0.000	0.117	0.204	0.243
Low		0.000	0.143	0.199
Mod			0.000	0.090
High				0.000

Supplementary Table 27. Lose Shift – Uncertainty level Komolgorov Smirnov p-values. The results of the two-sample Komolgorov-Smirnov (KS) hypothesis test comparing similarity between uncertainty level as reported by p-value.

Block	No	Low	Mod	High
No	1.000	0.177	0.001	<0.001
Low		1.000	0.051	0.006
Mod			1.000	0.585

High	1.000

Supplementary Table 28. Lose Shift – Uncertainty level mean and standard deviation (in proportion of shift trials after a losing trial).

Uncertainty	Mean	SD
No	0.752	0.195
Low	0.740	0.180
Mod	0.698	0.168
High	0.676	0.163

Supplementary Table 29. Probability Matching (all trials) – Blockwise t-test statistics. The results of the one-sample t-test comparing similarity between proportion of feature selection and ground truth as reported by t statistic, degrees of freedom, mean, standard deviation, and p-value.

Rule	Block	T-stat	DF	Mean	SD	P-value
Top	1	-12.279	59	0.673	0.206	<0.001
Top	2	-8.765	59	0.710	0.168	<0.001
Top	3	-10.339	59	0.574	0.132	<0.001
Top	4	-8.044	59	0.452	0.143	<0.001
Top	5	-7.588	59	0.568	0.186	<0.001
Top	6	-9.060	59	0.684	0.185	<0.001
Top	7	-7.800	59	0.794	0.204	<0.001
Top	8	-7.183	57	0.606	0.153	<0.001
Top	9	-4.860	58	0.491	0.172	<0.001
Top	10	-8.065	57	0.818	0.172	<0.001
Top	11	-5.340	56	0.773	0.180	<0.001
Mid	1	11.025	59	0.159	0.112	<0.001
Mid	2	6.236	59	0.140	0.086	<0.001
Mid	3	3.398	59	0.235	0.080	0.001
Mid	4	1.222	59	0.316	0.102	0.227
Mid	5	3.224	59	0.251	0.123	0.002
Mid	6	6.888	59	0.175	0.118	<0.001
Mid	7	6.954	59	0.105	0.117	<0.001
Mid	8	2.898	57	0.239	0.102	0.005
Mid	9	0.006	58	0.300	0.155	0.995
Mid	10	9.239	57	0.078	0.064	<0.001
Mid	11	3.734	56	0.111	0.082	<0.001
Bottom	1	10.220	59	0.168	0.127	<0.001
Bottom	2	9.709	59	0.150	0.096	<0.001
Bottom	3	12.229	59	0.191	0.090	<0.001
Bottom	4	9.350	59	0.232	0.110	<0.001
Bottom	5	8.508	59	0.181	0.119	<0.001
Bottom	6	8.504	59	0.142	0.102	<0.001

Bottom	7	5.488	59	0.100	0.142	<0.001
Bottom	8	9.112	57	0.155	0.088	<0.001
Bottom	9	5.741	58	0.209	0.146	<0.001
Bottom	10	6.173	57	0.105	0.129	<0.001
Bottom	11	4.535	56	0.117	0.144	<0.001

Supplementary Table 30. Probability Matching (all trials) – Uncertainty level t-test statistics. The results of the one-sample t-test comparing similarity between proportion of feature selection and ground truth as reported by t statistic, degrees of freedom, mean, standard deviation, and p-value.

Rule	Uncertainty	T-stat	DF	Mean	SD	P-value
Top	No	-15.618	177	0.761	0.204	<0.001
Top	Low	-13.168	176	0.721	0.180	<0.001
Top	Mod	-14.117	177	0.582	0.159	<0.001
Top	High	-8.856	118	0.471	0.159	<0.001
Mid	No	14.422	177	0.114	0.106	<0.001
Mid	Low	9.609	76	0.142	0.100	<0.001
Mid	Mod	5.413	177	0.242	0.103	<0.001
Mid	High	0.684	118	0.308	0.130	0.496
Bottom	No	12.240	177	0.124	0.136	<0.001
Bottom	Low	12.251	176	0.136	0.116	<0.001
Bottom	Mod	16.708	177	0.176	0.101	<0.001
Bottom	High	10.224	118	0.221	0.129	<0.001

Supplementary Table 31. Probability Matching (last 10 trials of block) – Blockwise t-test statistics. The results of the one-sample t-test comparing similarity between proportion of feature selection and ground truth as reported by t statistic, degrees of freedom, mean, standard deviation, and p-value.

Rule	Block	T-stat	DF	Mean	SD	P-value
Top	1	-13.559	48	0.940	0.031	<0.001
Top	2	-0.255	51	0.897	0.075	0.800
Top	3	3.350	44	0.813	0.126	0.002
Top	4	-2.759	53	0.529	0.189	0.008
Top	5	2.430	46	0.805	0.154	0.019
Top	6	0.098	49	0.901	0.083	0.922
Top	7	-6.690	50	0.975	0.027	<0.001
Top	8	3.250	44	0.820	0.144	0.002
Top	9	0.721	53	0.623	0.238	0.474
Top	10	-6.579	52	0.973	0.030	<0.001
Top	11	2.152	51	0.920	0.068	0.036
Mid	1	8.013	48	0.028	0.025	<0.001
Mid	2	-2.908	51	0.050	0.049	0.005
Mid	3	-5.148	44	0.108	0.120	<0.001
Mid	4	-0.181	53	0.295	0.186	0.857

Mid	5	-5.263	46	0.117	0.109	<0.001
Mid	6	-2.724	49	0.049	0.054	0.009
Mid	7	5.871	50	0.018	0.022	<0.001
Mid	8	-5.813	44	0.098	0.118	<0.001
Mid	9	-1.305	53	0.259	0.233	0.198
Mid	10	4.314	52	0.011	0.018	<0.001
Mid	11	-6.168	51	0.035	0.041	<0.001
Bottom	1	7.340	48	0.032	0.031	<0.001
Bottom	2	3.739	51	0.052	0.043	<0.001
Bottom	3	3.181	44	0.080	0.062	0.003
Bottom	4	4.415	53	0.176	0.126	<0.001
Bottom	5	2.166	46	0.079	0.091	0.036
Bottom	6	2.115	49	0.049	0.065	0.040
Bottom	7	3.105	50	0.007	0.016	0.003
Bottom	8	2.067	44	0.082	0.105	0.045
Bottom	9	1.248	53	0.118	0.106	0.217
Bottom	10	4.602	52	0.017	0.027	<0.001
Bottom	11	2.147	51	0.045	0.051	0.037

Supplementary Table 32. Probability Matching (last 10 trials of block) – Uncertainty level t-test statistics. The results of the one-sample t-test comparing similarity between proportion of feature selection and ground truth as reported by t statistic, degrees of freedom, mean, standard deviation, and p-value.

Rule	Uncertainty	T-stat	DF	Mean	SD	P-value
Top	No	-13.813	152	0.963	0.033	<0.001
Top	Low	1.030	153	0.906	0.076	0.305
Top	Mod	5.167	136	0.812	0.141	<0.001
Top	High	-1.129	107	0.576	0.219	0.261
Mid	No	10.237	152	0.019	0.023	<0.001
Mid	Low	-6.487	153	0.045	0.048	<0.001
Mid	Mod	-9.412	136	0.108	0.115	<0.001
Mid	High	-1.133	107	0.277	0.211	0.260
Bottom	No	8.456	152	0.018	0.027	<0.001
Bottom	Low	4.418	153	0.049	0.053	<0.001
Bottom	Mod	4.050	136	0.080	0.087	<0.001
Bottom	High	4.079	107	0.147	0.119	<0.001