Adaptive Behavior

Social contracts and human–computer interaction with simulated adapting agents
Alexis B Craig, Derrik E Asher, Nicolas Oros, Alyssa A Brewer and Jeffrey L Krichmar

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

ISAB

International Society of Adaptive Behavior

Additional services and information for *Adaptive Behavior* can be found at:

Email Alerts: http://adb.sagepub.com/cgi/alerts

Subscriptions: http://adb.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://adb.sagepub.com/content/21/5/371.refs.html

>> Version of Record - Sep 10, 2013

OnlineFirst Version of Record - Jul 29, 2013

What is This?

Social contracts and human–computer interaction with simulated adapting agents

Alexis B Craig[1], Derrik E Asher[1], Nicolas Oros[1], Alyssa A Brewer[1] and Jeffrey L Krichmar[1,2]

Abstract
Game theory is commonly used to study social behavior in cooperative or competitive situations. One socioeconomic game, Stag Hunt, involves the trade-off between social and individual benefit by offering the option to hunt a low-payoff hare alone or a high-payoff stag cooperatively. Stag Hunt encourages the creation of social contracts as a result of the payoff matrix, which favors cooperation. By playing Stag Hunt with set-strategy computer agents, the social component is degraded because of the inability of subjects to dynamically affect the outcomes of iterated games, as would be the case when playing against another subject. However, playing with an adapting agent has the potential to evoke unique and complex reactions in subjects because of its ability to change its own strategy based on its experience over time, both within and between games. In the present study, 40 subjects played the iterated Stag Hunt with five agents differing in strategy: exclusive hare hunting, exclusive stag hunting, random, Win-Stay-Lose-Shift, and adapting. The results indicated that the adapting agent caused subjects to spend more time and effort in each game, exhibiting a more complicated path to their destination. This suggests that adapting agents exhibit behavior similar to human opponents, evoking more natural social responses in subjects.

Keywords
Human–computer interaction, adaptive agents, Actor–Critic, Stag Hunt, game theory

1 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 (D. 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.

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

[1]Department of Cognitive Sciences, University of California, Irvine, California, USA
[2]Department of Computer Science, University of California, Irvine, California, USA

Corresponding author:
Alexis Craig, Department of Cognitive Sciences, University of California, Irvine 2220; Social and Behavioral Sciences Gateway, Irvine, CA 92697-5100, USA.
Email: acraig1@uci.edu

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 synched up to the audio component. The computer opponents were programmed to cooperate in four out of six trials. While the speech-based and 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 rate 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 cooperate or compete. However, in the Ultimatum Game, two players must decide how to divide a sum of money amongst themselves. 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 were more likely to accept unfair behavior from a computer player rather than a human player (Rilling et al., 2004). This suggests that human subjects did not hold their computer opponents to the same social constructs they held other humans to, alluding to the issue of not considering the computer opponents used in these experiments as socially equal. 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 with both human and computer player conditions. 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 player conditions engaged 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, 2012; 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 either to 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, Zaldivar, 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 system, 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 (T4T) 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-specific tendencies and or phenotypes in social situations.

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.

The Stag Hunt has recently been used to test theory of mind assumptions, both through modeling and by human subjects against computer agents (Yoshida, Seymour, Friston, & Dolan, 2010; 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 investigating the formation of social contracts, decision-making, and theory of mind.

A major goal of the present study is to show that an agent with the ability to adapt to another player's gameplay more effectively challenges a subject. In many 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 cultivated through 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 (EQHare), exclusive stag hunting (EQStag), random, WSLS, 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.

2 Methods

2.1 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 were removed before analysis.
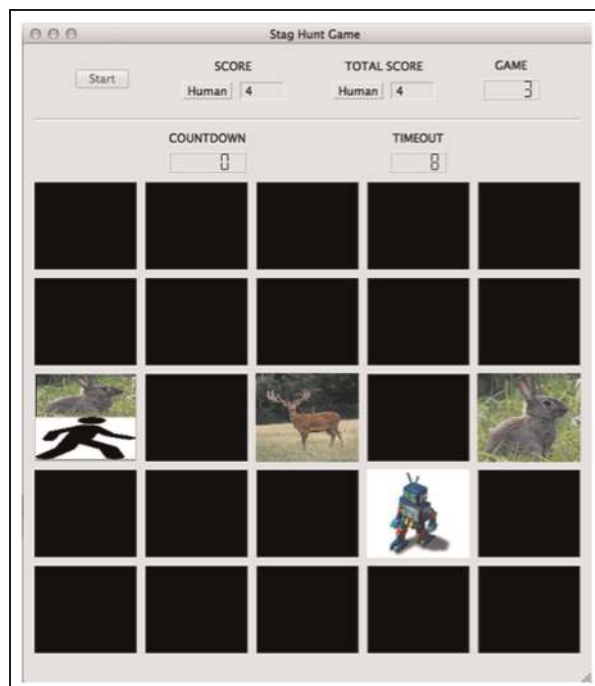
2.2 Computer interface for the Stag Hunt

Subjects played a variant of the Stag Hunt game against simulated agents, which was similar to the game used by Yoshida and colleagues (2010). This version of the Stag Hunt game differed from the traditional
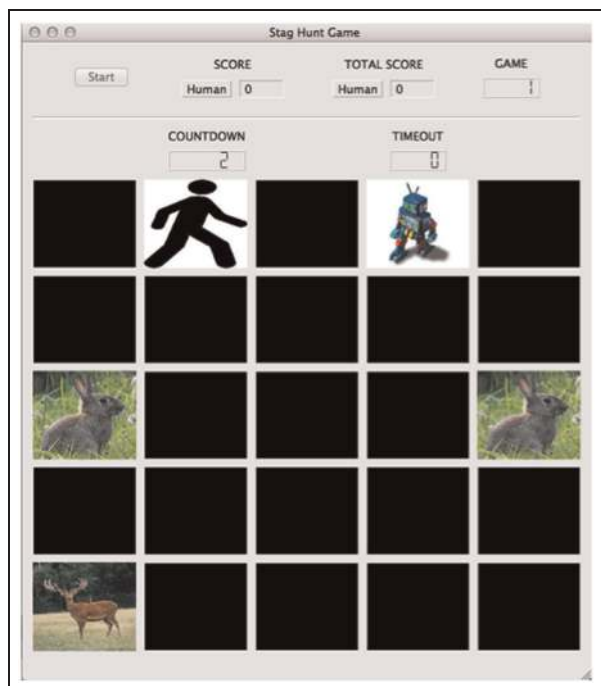
Figure 1. Screenshot of Stag Hunt game board. The game board included a 5 × 5 grid of spaces upon which the player (stick figure image), agent (robot image), stag (stag image), and hare (hare image) tokens resided. The screen included a button to start the experiment, the subject's score for the round, 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.

Figure 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.

version through the addition of spatial and temporal components 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.

The computer interface consisted of a 5×5 grid on which a stag token, two hare tokens, a subject token, and an agent token were placed (Figure 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.

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 each 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 ($\approx$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 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,
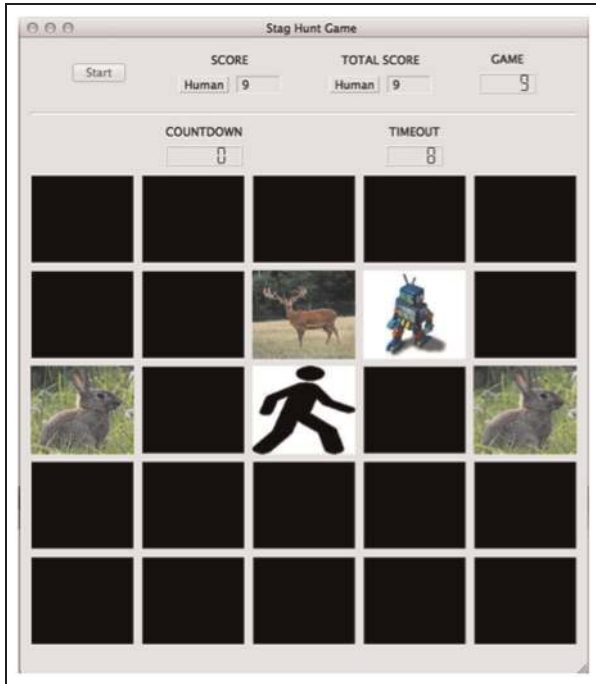
Figure 3. Screenshot of stag capture. In order to catch a stag, both the player and agent tokens needed to be in squares adjacent to the stag token, whether horizontally, vertically, or diagonally adjacent. Both the player and the agent required the intention of catching a stag. It was not sufficient 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.

horizontally, or diagonally (Figure 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). During the games, the subjects saw their total scores for the current round as well as a 10-second countdown timer 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 subject was 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.

2.3 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

Table 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

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 is 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, to give 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 (Table 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 (Table 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:

$$\delta(t) = r(t) + V(s, t) - V(s, t-1)) \qquad (1)$$

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 time-step. $r(t)$ for the Reward Critic was given as:

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

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

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

The Critic's state table was updated by:

$$V(s, t + 1) = V(s, t) + \delta(t) \quad (4)$$

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, 6, and 7. Equation 5 is given for the condition in which the Adapt agent hunted a hare.

$$V(hare, s, t + 1)) = V(hare,\ s, t) + 1 - p[hare]*\delta(t)$$
$$V(stag, s, t + 1)) = V(stag,\ s, t) - p[stag]*\delta(t) \quad (5)$$

$V(hare, s, t)$ 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.

$$V(stag, s, t + 1)) = V(stag,\ s, t) + 1 - p[stag]*\delta(t)$$
$$V(hare, s, t + 1)) = V(hare,\ s, t) - p[hare]*\delta(t) \quad (6)$$

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:

$$p[hare] = \frac{e^{V(hare, s, t)}}{e^{V(hare, s, t)} + e^{V(stag, s, t)}}$$
$$p[stag] = 1 - p[hare] \quad (7)$$

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.

2.4 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 10-second timer to run out during a game. End of experiment payments ranged from US$10 to US$21.

3 Results

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 4a) or if the other player was further away from a hare (Figure 4b). Figures 4(c) and 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.

Subject performance varied depending on the type of agent played (Figure 5; Table 2). In all cases, the
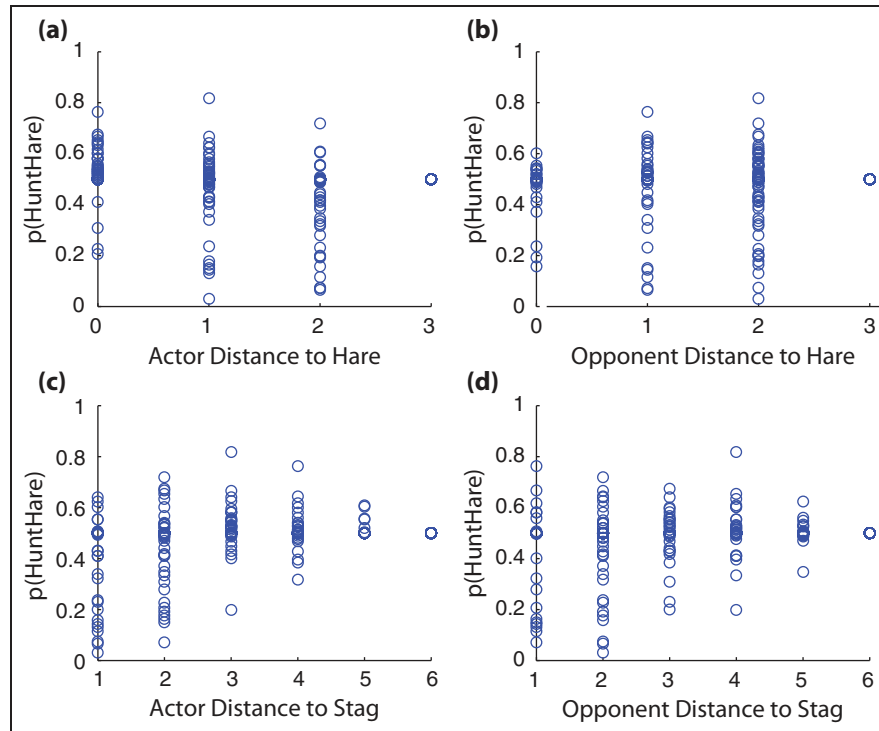
Figure 4. Scatter plot of Actor state table data. Data for all subjects were taken from the Actor state tables of the Actor–Critic models used in the Stag Hunt experiment. (a) shows the probability for the adapting agent to hunt the closest hare from each possible distance to closest hare target, while (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.
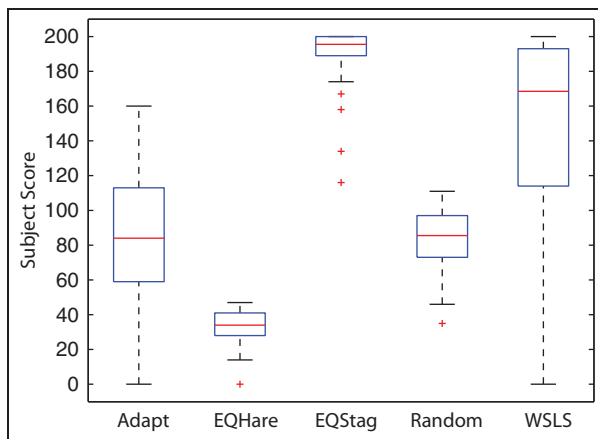

Figure 5. Subject scores against agent strategy. For each boxplot, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually as '+' symbols. The data were not 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 2).

Table 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					$< .0001$

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
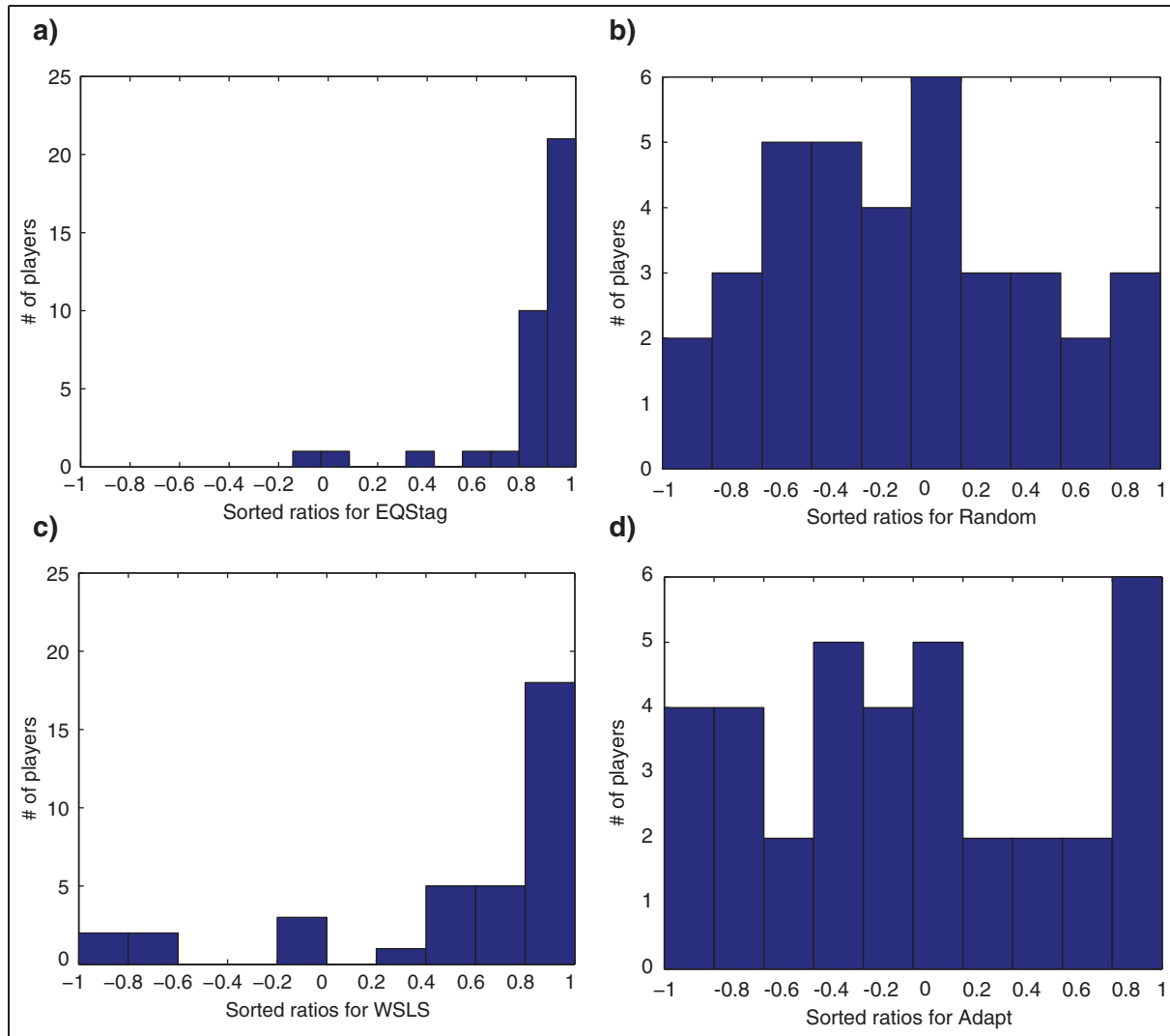
Figure 6. Ratio of stag to hare catches for all participants. The ratio of stag to hare catches was calculated by the equation $q_{hare:stag} = (n_{stag} - n_{hare})/(n_{stag} + n_{hare})$, in which $q_{hare:stag}$ is the ratio for a given subject, n_{stag} is the number of stags captured during a given 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.

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.

To understand how individual subjects altered their strategy when playing with different agents, we calculated the normalized ratios of stag to hare catches for each subject in each condition (Figure 6). The ratio was calculated by using the equation,

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

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 EQHare and positive one representing EQStag. In order to show the distribution of hunt behavior in subjects, Figure 6 shows how the subject hunted with an EQStag, Random, WSLS, and Adapt 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

**Table 3.** Color-coded chart of equilibrium alignment for individual subjects against each agent strategy.

Sub #	Adapt	EQStag	Random	WSLS	Avg	Key
11						: 0.9 < x
12						: 0.7 < x < 0.9
13						: 0.5 < x < 0.7
14						: -0.5 < x < 0.5
15						: -0.5 > x > -0.7
16						: -0.7 > x > -0.9
18						: -0.9 > x
21						
22						
23						
24						
26						
28						
31						
33						
34						
35						
36						
37						
38						
41						
42						
43						
44						
45						
46						
47						
48						
51						
52						
53						
54						
55						
56						
57						
58						

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.

WSLS conditions. As expected, subjects showed a bias toward stag hunting when playing against EQStag (Figure 6a), which suggests that they found a high-payoff equilibrium. In Figure 6(b), subjects playing a Random agent had a somewhat normal distribution of hunting tendencies with the peak being a mixture of stag and hare hunting. In Figure 6(c), subjects playing a WSLS agent had a bias toward stag hunting, as was also seen in EQStag, and was likely a result of the high-payoff equilibrium. In Figure 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 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 again 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
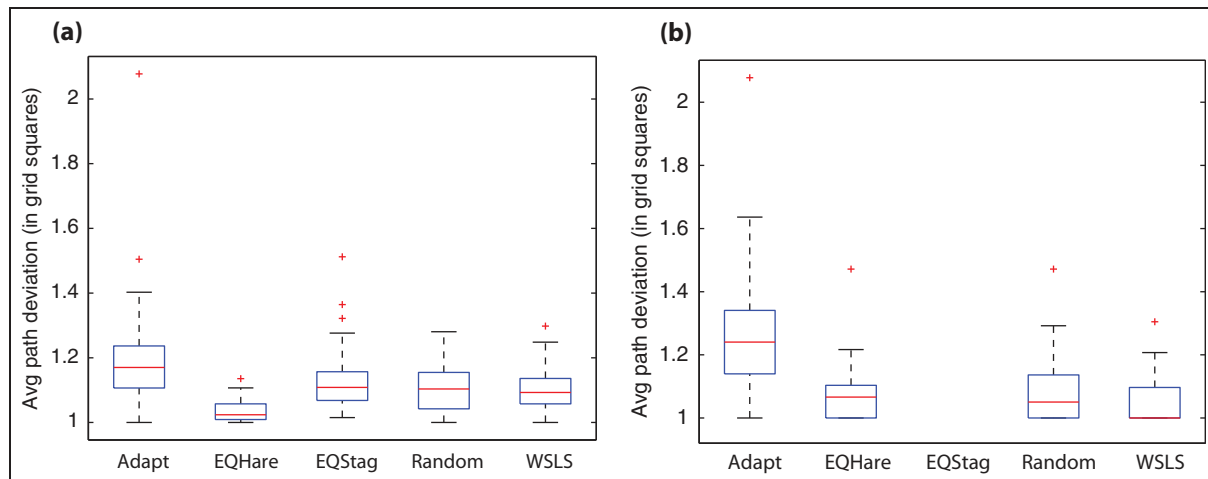
Figure 7. Average path deviation ratio over all subjects for each agent strategy. The boxplots have the same notation as in Figure 5. The data were not normally distributed; therefore subject performance against different agents was compared using Wilcoxon rank-sum tests (Bonferroni corrected, $p < .005$ 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 5).

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.

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 were analyzed over all outcomes (Figure 7a; Table 4), and also specifically games in which the subject failed to catch either a stag or a hare (Figure 7b; Table 5). 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 of each subject in each game) and subtracting that number from the subject's total distance traveled in each game (calculated by summing the distances between each move). 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

Table 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

Table 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

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 7b; Table 5), 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
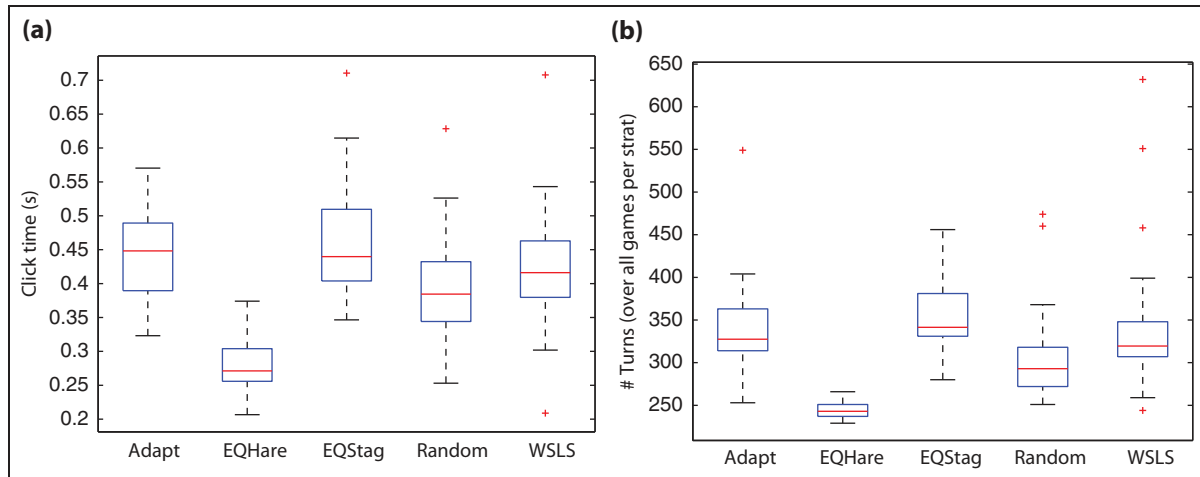
Figure 8. (a) The average time between mouse clicks for subjects during agent strategy: Adapt, EQHare, EQStag, Random, and Win–Stay–Lose–Shift (WSLS). The data were not normally distributed; therefore subject performance against different agents was compared using Wilcoxon rank-sum tests (Bonferroni 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.

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 was to time out. All timeout data were 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 with the human's average path deviation in Adapt, $\sim$1.2). For reference, each of the other agent types had an average path deviation of 1 (direct path) due to their inability to switch targets mid-game. In addition to using path deviation to give insight into the player's intention, move data in the Adapt condition were analyzed to see which player arrived at the stag first. Subjects arrived first at the stag 36% ($\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.

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 8; Tables 6 and 7). 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 with the Random condition (Figure 8a; Table 6).

Table 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

Table 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

Subjects took nearly significantly longer between mouse clicks when playing with the Adapt agent compared with 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 8b; Table 7). EQHare was found to have significantly fewer turns taken when compared with 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.

4 Discussion

Economic game theory has had a long, productive history of predicting and describing human behavior in cooperative and competitive situations (Maynard Smith, 1982; Nowak, Page, & Sigmund, 2000; Skyrms, 2001). The theory of games has also been used to illuminate the neural basis of economic and social decision-making (M.D. 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 8(b), subjects took significantly more turns when playing with the Adapt agent than the Random agent. This could indicate either that 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 were 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 support for 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 8a), 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 with

other conditions. In Figure 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 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 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 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 still comparable average value to the average human path deviation. The Adapt agent's path deviation behavior indicates that it was 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 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.

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 6d). About half of the subjects in the Adapt data 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 6d). In contrast, the ratio of stag-to-hare catching against Random agents was somewhat normally distributed with a peak towards equal stag and hare hunting (Figure 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 6(a) shows the EQStag data and Figure 6(c) shows the WSLS data. Both EQStag and WSLS appear to be heavily biased towards EQStag. In the case of EQStag, stag hunting was obviously encouraged by the fact that the agent hunted 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 was also hunting stags, which subjects playing to maximize their score should have done as predicted by game theory. Accompanying these histograms, the equilibrium table (Table 3) shows each individual subject's personal bias in hunting over those four conditions, implying that many subjects had 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 3).

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 study regarding HCI/HRI in the game of Hawk–Dove using an adaptive model (Asher, Zhang, et al., 2012). The conclusions drawn from their 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, 2003; Caspi, Hariri, Holmes, Uher, & Moffitt, 2010; Hariri, 2002; Lesch et al., 1996; Young et al., 2007). Subjects carrying the short allele variant of 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. Thus, 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 that 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 variation in 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.

5 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).

Acknowledgements

Funding

References

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.

Asher, D. E., Zaldivar, A., 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.

Bevilacqua, L., & Goldman, D. (2011). Genetics of emotion. *Trends in Cognitive Sciences, 15*(9), 401–408.

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.

Caspi, A. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science, 301*(5631), 386–389.

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.

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.

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

Hariri, A. R. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science, 297*(5580), 400–403.

Homberg, J. R., & Lesch, K.-P. (2011). Looking on the bright side of serotonin transporter gene variation. *Biological Psychiatry, 69*(6), 513–519.

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.

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.

Krichmar, J. L. (2008). The neuromodulatory system: A framework for survival and adaptive behavior in a challenging world. *Adaptive Behavior, 16*(6), 385–399.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nature Neuroscience, 11*(4), 404–409.

Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review, 15*(1), 1–15.

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.

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.

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.

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.

Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science, 289*(5485), 1773–1775.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology, 62*(1), 23–48.

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.

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/

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

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

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

Thomaz, A. L., & Breazeal, C. (2008). Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence, 172*(6-7), 716–737.

Valluri, A. (2006). Learning and Cooperation in Sequential Games. *Adaptive Behavior, 14*(3), 195–209.

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.

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.

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. (T. Behrens, Ed.) *PLoS Computational Biology, 4*(12), e1000254.

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.

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

About the Authors

Alexis B Craig received a BS in cognitive science from the University of California, Los Angeles in 2011 and is currently pursuing her PhD in psychology–cognitive neuroscience at the University of California, Irvine. Her research interests include cognitive robotics, human–robot interaction, game theory, and adaptive technology.

Derrik E Asher received a BS in chemical physics from the University of California, San Diego in 2006, then went on to complete an MS in cognitive neuroscience at the University of California, Irvine in 2010 for work with computational models of neuromodulation, decision-making and human–robot interaction (HRI). He is currently working with genetic algorithms to evolve biologically plausible artificial neural network models that perform sensorimotor integration and transformations in order to achieve behavioral error minimization.

Nicolas Oros received a BSc, an MSc and a PhD in computer science and artificial intelligence at the University of Hertfordshire, Hatfield, UK in 2005, 2006, and 2010, respectively. In 2010, he joined the cognitive anteater robotics laboratory at the University of California, Irvine, as a postdoctoral research scholar. His research interests include computational neuroscience, artificial intelligence, swarm intelligence and robotics, artificial life, and evolutionary computation.

Alyssa A Brewer received a BS in biological sciences with departmental honors and an AB in comparative literature with interdepartmental honors in the humanities in 1996 from Stanford University. She then completed a dual graduate degree program, receiving her PhD in neuroscience in 2005 from Stanford University and her MD in 2007 from Stanford University school of medicine. Since 2007, Dr Brewer has been an assistant professor in the department of cognitive sciences at the University of California, Irvine. Her research focuses on sensory neuroscience, using behavioral, genetic, and high-resolution neuroimaging techniques to investigate questions ranging from the fundamental organization of human visual cortex, functional plasticity in visuomotor regions, and visual changes in dementia, to human–robot social interaction, decision-making, and the organization of human auditory cortex.

Jeffrey L Krichmar received a BS in computer science in 1983 from the University of Massachusetts at Amherst, an MS in computer science from The George Washington University in 1991, and a PhD in computational sciences and informatics from George Mason University in 1997. He spent 15 years as a software engineer on projects ranging from the PATRIOT Missile System at the Raytheon Corporation to air traffic control for the federal systems division of IBM. In 1997, he became an assistant professor at The Krasnow Institute for Advanced Study at George Mason University. From 1999 to 2007, he was a senior fellow in theoretical neurobiology at The Neurosciences Institute. He is currently an associate professor in the department of cognitive sciences and the department of computer science at the University of California, Irvine. His research interests include neurorobotics, embodied cognition, biologically plausible models of learning and memory, and the effect of neural architecture on neural function.