You Can't Gamble on Others: Dissociable Systems for Strategic Uncertainty and Risk in the Brain

W. Gavin Ekins[1], Ricardo Caceda, C. Monica Capra[1], and Gregory S. Berns[1*]

[1]Center for Neuropolicy and Economics Department, Emory University, Atlanta, GA 30322 USA

*Correspondance: gberns@emory.edu

Abstract

This paper tests whether strategic uncertainty employs circuits in the brain that encode risk and utility, or circuits that are involved in Theory of Mind (ToM). We compare participants' decisions in a stag-hunt game with an equivalent choice between Bernoulli lotteries where the probabilities are equal to the mixed Nash equilibrium of the stag hunt game. Behavioral data suggests that most participants are more willing to choose the payoff-dominant option in a stag-hunt game than the equivalent lottery. Neuroimaging shows activations in the regions of the brain traditionally associated with ToM are correlated with a participant's propensity to choose payoff dominant while activations in region associated with encoding utility are correlated with a participant's propensity to choose the risk dominant. This suggests that individuals who mentalize the other person are more likely to be cooperative than those who do not.

Introduction:

The gains from cooperation are immense, but individuals often fail to take advantage of opportunities that require cooperation, even when it is the best option for both parties involved. Why some individuals are able to profit from these activities while others do not has been a riddle for economists. Although both laboratory and field experiments have been conducted with cooperative games, the behavior of participants does not reveal the motivations behind the cooperation. By using brain imaging to open the "black box" of the brain, we can further our understanding of cooperation and, more importantly, why it fails.

Rousseau's stag hunt story warns us that cooperation can fail, even when it is in our best interest. In his story, two hunters set out to hunt a deer. Each man must take and keep a post in order to capture a deer. While one hunter is manning his post, a hare passes in front of him. At this point the hunter is faced with a decision: continue to man his post or attempt to chase the hare. Both hunters would rather have the deer than the hare; the deer provides the greatest amount of food for both hunters. The only reason for the hunter to chase the hare is if he believes the other has abandoned his post in an attempt to capture smaller game. In other words, if one hunter fears that the other will chase a hare that runs across his path, then it is in his best interest to also chase a hare (Skyrms 2001).

What is puzzling in Rousseau's story, which is the inspiration behind the stag hunt game, is there is no payoff incentive for either party to be uncooperative. Unlike the prisoner's dilemma game, in which it is profitable to be uncooperative; uncooperative behavior in the stag hunt game is a consequence of mistrusting the other player's motivations. Cooperating in the hunt for the stag has the greatest payoff, both socially and individually; but the mistrust of each other's motives leads the players to forgo the optimal outcome in return for a sure outcome. This kind of uncertainty of the other player's choice is considered strategic uncertainty, and has been a topic of interest to theorists and experimentalists alike (Harsanyi & Selten 1988; Van Huyck et al. 1997).

In game theory, the stag hunt game is a simultaneous game with complete information that can be expressed in matrix form (figure 1a), where (stag,stag) and (hare,hare) are both Nash equilibria in pure strategies. One of these equilibria (stag, stag) is preferred or payoff dominant, which means that both players are better off if this equilibrium is reached as compared with the alternative. In addition, there is an equilibrium in mixed strategies. Behaviorally, since the game has multiple solutions, and the payoff structure of the game does not provide a clear prediction of players' behavior, game theorists have resorted to a mixed Nash equilibrium to resolve the ambiguity. However, both economists and game theorists argue that the mixed Nash equilibrium's assumptions of human behavior are impractical due to its reliance on individuals' ability to randomize (Radner & Rosenthal 1982; Aumann 1987; Van Huyck et al. 1997; Rubinstein 1991). With respect to the pure strategy equilibria, one would expect that rational players would choose the payoff dominant equilibrium (stag, stag), yet this choice is risky. Indeed, Harsanyi and Selten (1988) introduced an alternative solution, the risk dominant, which in the context of the above game would be (hare, hare).

The payoff dominant equilibrium expresses a player's desire to maximize his payoffs. Since (stag, stag) has the greatest payoff, players who seek to maximize their own payoff will attempt to coordinate on the (stag, stag) Nash equilibrium. The risk dominant equilibrium expresses a player's desire to play it safe or minimize the cost of failing to coordinate on one of the two Nash equilibria. Consider the game in Figure 1. The payoffs are such that $A>C\geq D>B$ and $a>c\geq d>b$. If the column player unilaterally chose to move from (stag, stag), the row player would pay a greater opportunity cost than if he were to unilaterally move from (hare, hare). As such, the cost of failing to coordinate on (stag, stag) is greater than the cost for failure to coordinate on (hare, hare). If a player is unsure of the other player's motives, he risks less by choosing hare than by choosing stag. Therefore, the hare option is the risk dominant equilibrium.

The choice of risk dominance over payoff dominance is attributed to strategic uncertainty, but it is unclear if strategic uncertainty is a matter of risk preferences or strategic mistrust of the other player's actions. Straub (1995) has argued the strategic uncertainty is strictly different from risk preferences, but it is possible that players convert the uncertainty of the other player's decision into a game against nature or a lottery. If this is the case, the mechanisms used by individuals to evaluate random events can be applied to games with strategic uncertainty.

Although there are several experimental studies of the stag hunt game (Van Huyck et al. 1997; Straub 1995; Rankin et al. 2000; Battalio 2001) behavioral data cannot determine which mechanism individuals use to evaluate the game. In this situation, neuroscience can help resolve the question. By observing which parts of the brain are active during a decision, we can determine which mechanism an individual employs to make this type of decision.

Neuroscientists have argued that there are innate cognitive mechanisms that facilitate prediction of others' actions (Baron□Cohen 1989; Corcoran et al. 1995) Commonly known as the theory of mind (ToM), these mechanisms help individuals to create a mental model of another person. With this mental model, an individual can perceive a situation from the point-of-view of another person, whether she is involved in a strategic game or a false belief task. False belief tasks and economic game experiments have shown that several regions of the brain are commonly involved in mentalizing another person. These areas include the medial prefrontal cortex (MPFC), superior temporal sulcus (STS), temporal poles, fusiform gyrus, and both the anterior cingulate cortex (ACC) and the posterior cingulate cortex (PCC) (Vogeley et al. 2001; McCabe et al. 2001; Gallagher & Frith 2003; Rilling et al. 2004; Völlm et al. 2006; Tomlin et al. 2006) Autism research has also implicated the fusiform gyrus, in particular the fusiform face area (FFA), as part of the system used to mentalize others (Schultz et al. 2003; Dalton et al. 2005; Gobbini et al. 2011).

Other studies have shown that the encoding of risk and utility is regional as well. The ventral striatum has been shown to encode the magnitude of a monetary reward in both random and non-random outcome experiments (Breiter et al. 2001; B Knutson et al. 2001; O'Doherty et al. 2004; Abler et al. 2006). When varying the probability of an outcome, the ventral striatum also encodes the likelihood of a particular outcome occurring (Kuhnen & Knutson 2005; Dreher et al. 2006; Abler et al. 2006). These results are generalizable to primary rewards, such as food (Pagnoni et al. 2002; McClure et al. 2003). Similarly, the VMPFC has been implicated in encoding the magnitude of random events when a reward was given, but was unexpected (B Knutson et al. 2001; Ramnani and Owen 2004).

From these prior neuroimaging studies, we hypothesize that if a player of the stag-hunt game is using a mental model of the other person to make her decision, then the ToM regions should be active while playing the game. On the other hand, if she processes strategic uncertainty as random events, then we hypothesize that circuits in the brain that are traditionally associated with evaluating risk and utility will be active (ventral striatum and VMPFC).. To determine which mechanisms are employed in the stag-hunt, we designed an experiment which compares decisions made in the stag-hunt game with those made in an equivalent gamble. Using functional magnetic resonance imaging (fMRI) we tested if the ToM regions are more active in the stag-hunt game than in an equivalent gamble, or whether risk and utility regions are more active.

Methods:

Game:

To study how the brain processes strategic uncertainty and risk[1], we developed an experiment where participants make decisions in strategic games and comparable gambles. In the stag-hunt game, players are faced with a payoff dominant choice and risk dominant choice. In the payoff dominant choice, the player can maximize his earnings if their partner coordinates with them, but they receive the lowest payoff otherwise. In the risk dominant choice, the player receives less than the maximum of the payoff dominant choice, but receives more than the worst outcome regardless of their partner's decision. If the player trusts their partner to coordinate with them, then her best option is to choose the payoff dominant option. If she is unsure of the partner, then her best option is to choose the risk dominant option. Unlike the prisoner's dilemma game, there is no direct incentive not to choose payoff dominant. Thus, the choice of risk dominant must be related to a player's uncertainty of the other player's choice.

In these gambles the players faced the exact same payoffs as in the stag-hunt game, but the outcome of their decisions was determined by a random draw, not another person's choice.

Treatments:

We used a 2 x 2 factorial design. In the first factor, the participant was either playing a strategic game (S) or a lottery (L) with the same payoffs for each possible outcome. In strategic treatments, the participants were shown a standard normal-form game matrix (figure 1b). In the matrix, the participant's payoffs were the values to the left of the comma within each cell (A), while the other player's value was to the right of

[1] Risk as defined in Knight (1921)

the comma (a). The participant was always player 1 and had the choice of Up or Down. The player 2's decisions were registered by another participant before the experiment began.

In the lottery treatments, participants could choose between two lotteries, rather than playing a strategic game. The lotteries were displayed as a matrix with an up or down decision (figure 1b). The Up decision was a lottery with outcomes A or B and the Down decision was a lottery with outcomes C or D. The probability of each outcome was displayed at the top of the columns. These probabilities were based on the mixed Nash equilibrium probabilities from the equivalent strategic-game matrix. The outcome of the chosen lottery was determined by dice at the end of the experiment.

Figure 1

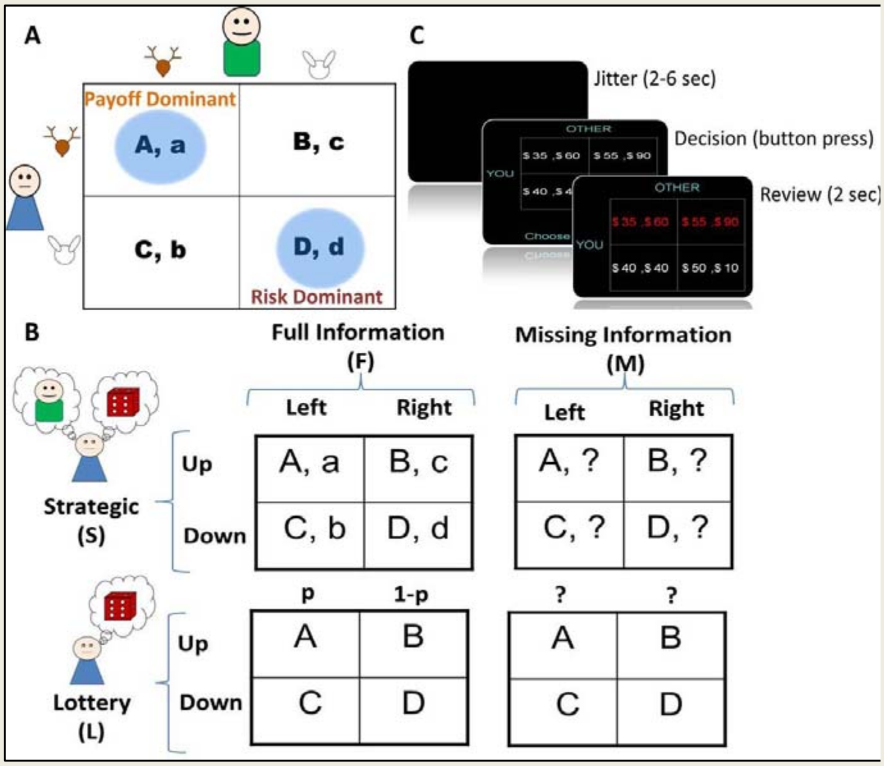

Figure 1. A: Standard normal-form stag-hunt game where A>C≥D>B and a>c≥d>b. Blue circles indicate the Nash equilibria. **B:** Game matrices presented in each treatment. The probabilities (p) in lottery treatment are the mixed Nash equilibrium probabilities from the equivalent strategic treatment. Note that the strategic treatment can invoke ToM networks, risk and utility circuits, or both; while the lottery only invokes the risk and utility circuits. **C:** Screen shots of game played in scanner. Jitter was a blank screen present for 2-6 seconds uniformly distributed. Decision screen lasted until the subject responded. The review screen, which highlights the decision, had a duration of 2 seconds.

The second factor was whether the participant was provided with all the information regarding the other player's payoffs. When there was full information (F), participants were given the payoffs of the other player in the strategic treatment and the probabilities of each outcome in the lottery treatment (Figure 1b). When there was missing information (M), a "?" was substituted for the payoffs of the other player in each cell during the strategic treatment or the probabilities at the top of the column in the lottery treatment.

In total there were 4 treatments: strategic game with full information (SxF), strategic game with missing information (SxM), lottery with full information (LxF), and lottery with missing information (LxM). Each treatment contained the same 56 game matrices. Within a treatment, we varied payoff incentives. That is, we tried to make the payoff-dominant and risk-dominant more attractive in some trials and less attractive in others. In addition, we included both symmetric and asymmetric payoffs, and varied matrices with respect to the mixed Nash equilibrium probabilities. We varied the matrices to ensure the robustness of our findings.

Procedures:

We recruited 30 participants from Emory University (M: 12 F: 18 age: 18-58). Each participant provided written consent and was given written instructions with a verbal overview. Participants had an opportunity to practice the task with the input device[2] while in the scanner.

Following the practice, participants completed 4 runs, each consisting of 56 trials. In each run, there were 8 blocks of 7 trials. All trials within a block had the same treatment. Each block of a particular treatment was repeated twice per run and each treatment block was separated with a block of another treatment. This ensured that the observations were balanced in each run. Each game matrix was repeated in each of the treatments, but each matrix was only displayed once per run.

In each trial, there were three phases. The first phase displayed the game matrix to the participant (Figure 1c). The participant was then given the option to choose up or down. Payoff dominant and risk dominant choices rotated such that the Up choice was not always payoff dominant and the Down choice was not always risk dominant. In the second phase, once the participant had made her decision, the row of her choice was highlighted for 2 seconds (Figure 1c). In the third phase, a blank screen was shown for a length of time between 2-6 seconds randomly chosen from a uniform distribution. Following the jitter screen, the next trial began.

Once all runs were completed, participants were randomly assigned one of the trials for which they were paid. Dice were used to randomly select the trial for which the participant is paid. If the trial was strategic, participant's choices were compared with a set of preregistered decisions and the participant was paid based on the matrix of both players' decisions. If the trial was a lottery, the participant rolled a die to determine which of the two outcomes they received.

fMRI data acquisition and analysis:

[2] Standard 4-button box

Functional imaging was performed with a Siemens 3 T Trio whole-body scanner. T1-weighted structural images (TR = 2600 ms, TE = 3.02 ms, flip angle = 8°, 240 × 256 matrix, 176 sagittal slices, 1 mm^3 voxel size) were acquired for each subject prior to the four experimental runs. For each experimental run, T2*-weighted images using an echo-planar imaging sequence were acquired, which show blood oxygen level-dependent (BOLD) responses (echo-planar imaging, TR = 2000 ms, TE = 30 ms, flip angle = 73°, FOV = 192 mm × 192 mm, 64 × 64 matrix, 33 3.5-mm thick axial slices, and 3 x 3 x 3.5 mm voxels).

fMRI data were analyzed using SPM8. Data were subjected to standard preprocessing, including motion correction, slice timing correction, normalization to an MNI template brain, and smoothing using an isotropic Gaussian kernel (full-width half-maximum = 8 mm). A standard 2-stage random-effects regression model was used for statistical inference.

To test the hypothesis that a strategically uncertain game employs ToM regions of the brain, we first separated the payoff and risk dominant decisions into different imaging models. By separating the decisions, we avoided the issue of missing observations in particular treatments. In the first imaging model, which we call the payoff dominant model, trials were categorized into separate conditions for each of the four treatments in which the participant made the payoff dominant decision. A nuisance regressor for all risk dominant decisions, regardless of treatment, was added to account for the remaining decisions.

In the second model, which we call the risk dominant model, was based on a similar structure as the first. Instead of categorizing trials when the participants chose payoff dominant, the trials were sorted into the four treatment conditions when the risk dominant choice was made. Similarly, the nuisance regressor in the second model was for all payoff dominant decisions regardless of treatment.

Each trial began when the participant was shown the game matrix and ended when they completed their decision. Task regressors were convolved with a standard canonical hemodynamic response function. Six motion regressors were added to compensate for participant movement: 3 position and 3 rotation variables. The task regressors, modulated regressors, and motion variables along with a constant were generated for each run. Runs with missing observations were removed from the model. None of the runs were removed from the payoff-dominant model, while 6 of the runs were removed from the risk-dominant model.

In order to absorb within-subject variation due to matrix-specific incentives, each condition was modulated[3] by a payoff-index. The index was based on the matrix payoffs, which reflected the relative incentives to choose the payoff dominant option. The index was a modified Nash product for a normal form game: —— . Since there were no payoffs for the other player in the lottery treatments, we removed the elements of the equation that are part of the other player's payoffs.

To test the hypothesis that ToM underlies coordination, we used predefined ROIs for the following regions: ACC, PCC, STS, temporal poles, mPFC, and fusiform gyrus. To test the alternative hypothesis that strategic uncertainty is processed similarly to risk, we used predefined ROIs in the ventral striatum and vmPFC. Regions were defined based on the Wake Forest University pickatlas. We extracted regression coefficients in each ROI for both the payoff dominant model and the risk dominant model.

Figure 2

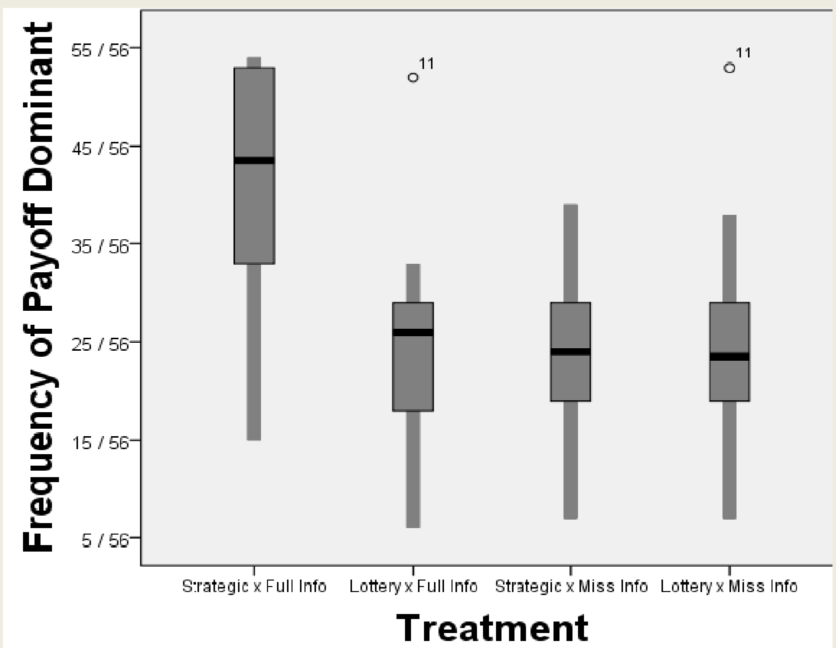

Figure 2. Box plot of the number of times payoff dominant is chosen by each participant by treatment. Sample size is 30 participants. T-test shows that LxF, SxM, and LxM treatments are statistically the same ($p<0.05$). SxF treatment is statistically different from all others ($p<0.05$).

Results:

We begin by examining participant's preference for the payoff dominant choice in the four treatments. On average, participants in the SxF treatment chose payoff dominant option 44 out of 56 different matrices. This was significantly larger than the other three treatments, which had averages almost half that of the SxF treatment (figure 2). LxF, SxM, and LxM were statistically indistinguishable from each other. Although most participants showed a stronger preference for the payoff dominant choice in the SxF treatment than the other, there were individuals who did not differentiate between treatments or differentiated differently across treatments.

[3] See Rajapakse et al. 1998 for explanation of statistical methodology.

The extent to which a participant differentiates between the treatments may be indicative of how they view the other player. Those who showed no behavioral change between the strategic and lottery treatments (i.e., the number of times they chose payoff dominant was the same in both treatments) may not have differentiated between the two types of decisions. If this were the case, it is likely that there would be no difference between how the brain processes strategic uncertainty in a game and a lottery, and the participant viewed the strategically uncertain game as a lottery. On the other hand, if an individual behaved markedly different in each treatment (i.e., there was a large difference between the strategic and lottery treatments in the number of times payoff dominant was chosen) then it is more likely there would be differences in brain activity between the two treatments, and the participant viewed the other person as something different than a lottery.

Imaging Results:

Table 1

Theory of Mind	PD	RD	Utility and Risk
Anterior Cingulate Cortex (ACC)	**0.44***	*-0.31*	
Posterior Cingulate Cortex (PCC)	**0.69*****	*0.13*	
Superior Temporal Sulcus (STS)	**0.40***	*-0.12*	
Temporal Poles	*0.28*	*-0.15*	
Medial Prefrontal Cortex (mPFC)	*0.23*	*-0.34*	
Fusiform	**0.64*****	*-0.02*	
	0.07	**-0.41***	Ventral Striatum
	0.33	*-0.34*	Ventromedial Prefrontal Cortex (vmPFC)

*p<0.05, **p<0.01, ***p<0.001

Table of Pearson correlations between a participant's change in activations for each region of interest (ROI) and the change in their behavior. The payoff dominant model (PF), on the left, correlates the change in activations for each region during a payoff dominant decision with the change in the number of times the participant chose the payoff dominant option. The risk dominant model (RD), on the right, correlates the change in activations for each region during a risk dominant decision with the change in the number of times a participant chose the risk dominant option. Change in activations were calculated by extracting the beta values from the ROI in the strategic with full information > lottery with full information (SxF-LxF) contrast. Change in behavior was calculated from the difference in the number of payoff dominant (risk dominant) decisions made in the strategic full information (SxF) treatment and those made in a lottery with full information (LxF) treatment.

Due to the lack of statistical difference in the behavioral data between the LxF, LxM, and SxM, we restricted our analysis of the brain images to the differences between SxF and LxF, which had significant behavioral differences. When we extracted the regression coefficients for the predefined ROIs, we found that, on average, none of the regions had significantly different activations between the two treatments for either the payoff- or risk-dominant models. However, as noted above, it is unlikely that all participants viewed the other player the same.

The heterogeneity of behavior between the participants could have been due to differences in brain mechanisms used during the task. For example, those who chose payoff dominant more often in the strategic than lottery treatments would be expected to show greater activity in strategic treatments when compared to participants who did not show any behavioral differences across these treatments. Any regions showing this pattern would be likely candidates for mediating the processing of strategic uncertainty differently than risk. Therefore, if the difference between activations in the strategic and lottery treatments in ToM regions is correlated with the difference between treatments in frequency of payoff dominant choices, i.e. the number of times payoff dominant is chosen, then it is likely that ToM regions are being employed when a payoff dominant decision is made.

When we examined the differences in number of times a participant chose the payoff dominant option in SxF and the LxF treatment in the payoff dominant model, we found that four of the seven ToM regions correlated with the behavior. Using Pearson correlations, we found that PCC and the fusiform gyrus were the most significantly correlated with behavior (Figure 4). The ACC and the STS were weakly, but significantly correlated as well. None of the risk and utility regions were correlated with the subjectwise differences in behavior in the payoff-dominant model.

We repeated this analysis for the risk dominant model, but instead of using the difference in payoff dominant choices, we used the differences in risk dominant choices between the SxF and LxF treatment. In this case, none of the ToM regions correlated with behavior, but one of the risk and utility regions, the ventral striatum, was significantly correlated with behavior (Figure 3).

We explored several alternative models. This included a model which pooled both payoff dominant and risk dominant decisions into the task variable. This model yielded significant activations, but failed to show significant correlations with behavior. We also developed a single model with a separate regressor for treatment and choice. For example, the SxF payoff dominant regressor and the SxF risk dominant regressor were used in the same model, but due to too many missing observations, this model had insufficient power. We explored all models without the index modulator, but none of the ROIs showed signficant effects, likely due to the variability of matrix specific incentives.

Figure 3

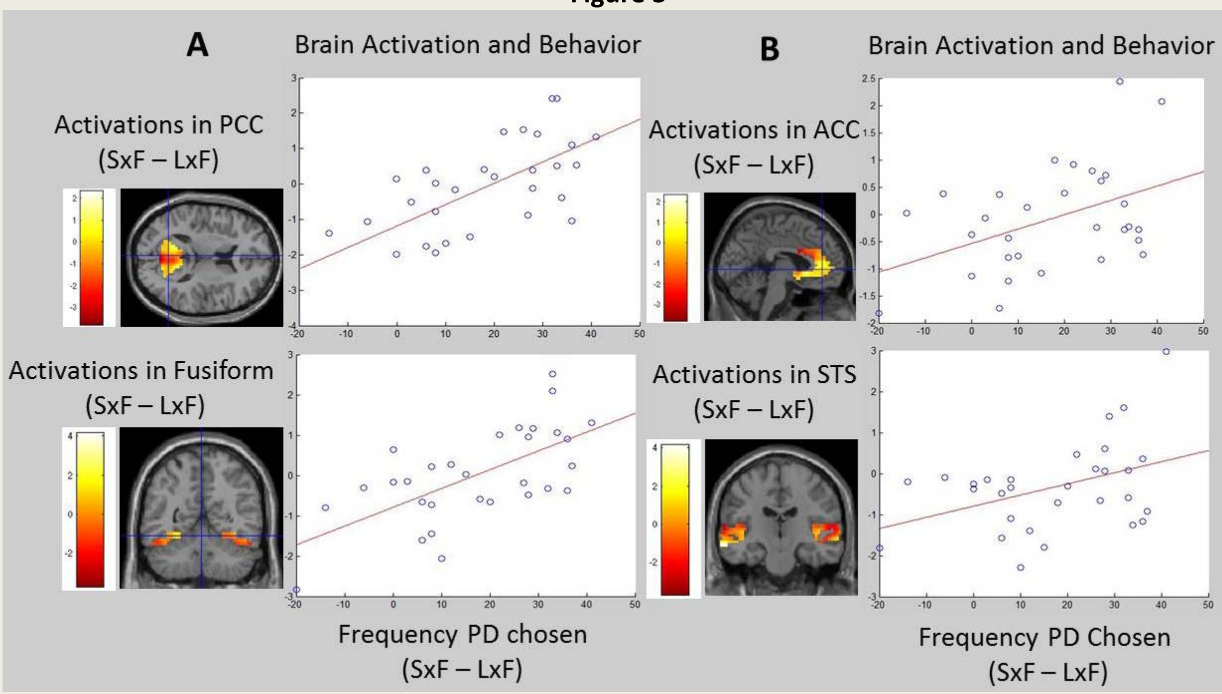

Figure 3. Scatter plots of change in behavior (x axis) and the change in activations (y axis). Change in behavior is the difference in the number of time payoff dominant (PD) is chosen in the strategic with full information (SxF) treatment and the lottery with full information (LxF) treatment. Change in activations are the regression coefficients from the ROI of the payoff dominant model in the strategic with full information > lottery with full information constrast (SxF-LxF). ROI mask used in extracting coefficients, to the left of scatter plot, with aggregate activation of all participants. **A:** Plots of posterior cingulate cortex (PCC), top left, and the fusiform, bottom left. Both curves had a Pearson correlation coefficient greater than 0.60 (see figure 3) and are significant ($p<0.001$). Correction for multiple comparisons (bonferroni, n=8) did not change the significance. **B:** Plots of anterior cingulate cortex (ACC), top right, and superior temperal sulcus (STS), bottom right. Both curves have a Pearson correlate greater than 0.40 (see figure 3) and are significant ($p<0.05$). When correcting for multiple comparisons (bonferroni, n=8), STS was no longer statistically significant.

Discussion:

The present study seeks to explain the motivations behind a risk dominant decision in a strategically uncertain game where the optimal outcome for both players is the payoff dominant choice. As other behavioral experiments have shown (Straub 1995), we found that participants played a mixture of risk and payoff dominant decisions in a stag-hunt game. In comparison, participants were more likely to choose risk dominant in a choice between Bernoulli lotteries with payoffs and probabilities equivalent to a stag-hunt game. Moreover, when the payoffs of the other player were hidden, participants were equally likely to choose the payoff dominant option as the lottery, which was significantly less than the stag-hunt with payoffs. Our behavioral data suggests that the choice of payoff dominant in a stag-hunt game is more than a decision of risk and utility. In addition, the payoffs of the other player are necessary to increase payoff dominant decisions, which suggest that participants are using the payoffs to generate a prediction of the other player's actions.

To help explain the behavioral differences between the stag-hunt game and the lottery, we examined the difference between fMRI activation during a stag-hunt game and lottery decision, both with full information. When participants made payoff dominant decisions, we found that activations in four of the six regions associated with Theory of Mind (ToM) functions were correlated with a participant's propensity to choose payoff dominant, while none of the risk and utility regions correlated with behavior. The increased activation in these four regions: ACC, PCC, STS, and fusiform; suggest that participants who mentalize the other player are more likely to attempt to coordinate on the higher-payoff, higher-risk equilibrium. When participants made a risk dominant decision, we found that none of the ToM regions were correlated with a participant's propensity to choose risk dominant, but activations in the ventral striatum, which is associated with encoding value, was correlated with this behavior.

Even though a majority of ToM regions were correlated with behavior, the temporal poles and MPRC were not. The lack of significance in the temporal poles, can be explained by the design of the experiment. Galloghar and Firth (2003) have argued that the temporal poles are primarily for recognition and recall of an individual. Since the experiment used an anonymous partner and provided no feedback after each trial, there was no need for recognition or recall. The lack of significance in the MPFC, which is the most consistently activated during ToM tasks (Carrington & Bailey 2009), is more puzzling. However, there is evidence that the MPFC is equally employed in a lottery (Wu et al. 2011).

Although we found significance in the STS for the payoff dominant model and the ventral striatum for the risk dominant model, they were relatively weak ($p<0.05$, uncorrected for multiple comparisons). When we corrected for multiple comparisons[4], these regions were no longer significant, but the ACC, PCC, and the fusiform gyrus remained significant in the payoff dominant model. Nevertheless, the disparate sorting of the regions into their hypothesized category at lower criteria for significance suggests the validity of the uncorrected tests, even for the STS and ventral striatum. For example, none of the ToM regions were significant in the risk dominant model, and none of the risk and utility regions were significant in the payoff dominant model. Moreover, all other treatment contrasts did not yield significant ($p<0.05$) correlations between activations and behavior for either model. Therefore, the failure of those regions to survive a multiple comparison test is more likely an effect of a limited sample size, rather than a false positive in the uncorrected tests.

Our initial hypothesis was that if individuals in strategically uncertain games see the other player as a person, ToM regions should be correlated with the presence of the other player. Although we found that behavior is correlated with activations in these regions, this connection only emerged when accounting for the heterogeneity of participants. The continuum of greater preferences for payoff dominant in a strategic game over a lottery suggests that it is specific to the individualand may not be generalizable to every person.

It is also possible that the increase in frequency of choosing payoff dominant was a reflection of increased activation in the ToM regions. Just as Knutson et al. (2008) found that stimulating the nucleus accumbens in males increased their preference for risk, stimulating the ToM regions my also increase an individual's propensity to choose the payoff dominant option. As such, activities and treatments that promote mentalizing the other player may help to increase cooperative behavior.

The findings that activation in ToM region was correlated with a propensity to choose payoff dominant have particular implications in economic development as well. Since the choice of risk dominant is both an individual and socially inferior solution, there are Pareto improvements when individuals prefer the choice of payoff dominant (Harsanyi & Selten 1988). The results of our experiment suggest that mentalizing the other player is an important part of the decision to choose the payoff dominant option. As such, activities which stimulate mentalizing others can lead to Pareto improvements in the economy as a whole.

---

[4] Bonferroni correction with n=8 for both the payoff dominant model and the risk dominant model. Significance threshold of 0.05 becomes 0.006.

An example of such Pareto improvements can be found in bargaining. In legal arbitration, all parties are better off if a deal can be struck outside the formal legal system. If one party mistrusts the other, then he is better off hiring a lawyer rather than risking inadequate representation. Much like the stag hunt, both parties would like to bargain outside of the legal system, but mistrust make bargaining outside the legal system difficult (Baird et al. 1998). Face to face meetings, which will help both parties mentalize the other, may be an important tool for facilitating an agreement outside the law. Therefore, requiring a plaintiff to have face to face meetings with defendants my reduce the need for trials and demand for court resources.

It is possible to argue from these findings that societies who view the economy as a set of personal interactions should have higher productivity and wealth than societies that view the economy as a set of gambles. Such an argument creates a conundrum: why do the formal, impersonal institutions of the West outperform the informal, personal institutions elsewhere (North 1981; Greif 2006)? The framework of a two person game limits the generalization to bilateral interaction and should not be applied to societal interaction. N-person stag hunt games, or threshold games, are more suitable for describing societal interaction. As such, future research of more social assurance games will help to resolve such a question.

Bibliography

Abler, Birgit, Henrik Walter, Susanne Erk, Hannes Kammerer, and Manfred Spitzer. 2006. “Prediction error as a linear function of reward probability is coded in human nucleus accumbens.” *NeuroImage* 31 (2) (June): 790-795. doi:10.1016/j.neuroimage.2006.01.001.

Anderson, Goeree, and Holt. 2002. “The Logit Equilibrium: A Perspective on Intuitive Behavioural Anomalies.” *Southern Economic Journal* 69 (1): 21-47. doi:urn:nbn:nl:ui:29-107479.

Baird, Douglas G., Robert H. Gertner, and Randal C. Picker. 1998. *Game theory and the law*. Harvard University Press, September 1.

Baron□Cohen, Simon. 1989. “The Autistic Child’s Theory of Mind: a Case of Specific Developmental Delay.” *Journal of Child Psychology and Psychiatry* 30 (2) (March 1): 285-297. doi:10.1111/j.1469-7610.1989.tb00241.x.

Battalio, Raymond, Larry Samuelson, and John Van Huyck. 2001. “Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games.” *Econometrica* 69 (3) (May 1): 749-764. doi:10.1111/1468-0262.00212.

Breiter, H C, I Aharon, D Kahneman, A Dale, and P Shizgal. 2001. “Functional imaging of neural responses to expectancy and experience of monetary gains and losses.” *Neuron* 30 (2) (May): 619-639.

Camelia Kuhnen, and Brian Knutson. 2005. *The Neural Basis of Financial Risk Taking*. EconWPA, September. http://ideas.repec.org/p/wpa/wuwpex/0509001.html.

Carrington, Sarah J, and Anthony J Bailey. 2009. “Are there theory of mind regions in the brain? A review of the neuroimaging literature.” *Human Brain Mapping* 30 (8) (August 1): 2313-2335. doi:10.1002/hbm.20671.

Corcoran, Rhiannon, Gavin Mercer, and Christopher D. Frith. 1995. “Schizophrenia, symptomatology and social inference: Investigating ‘theory of mind’ in people with schizophrenia.” *Schizophrenia Research* 17 (1) (September): 5-13. doi:10.1016/0920-9964(95)00024-G.

Dalton, Kim M, Brendon M Nacewicz, Tom Johnstone, Hillary S Schaefer, Morton Ann Gernsbacher, H H Goldsmith, Andrew L Alexander, and Richard J Davidson. 2005. “Gaze fixation and the neural circuitry of face processing in autism.” *Nat Neurosci* 8 (4) (April): 519-526. doi:10.1038/nn1421.

Dreher, Jean-Claude, Philip Kohn, and Karen Faith Berman. 2006. “Neural Coding of Distinct Statistical Properties of Reward Information in Humans.” *Cerebral Cortex* 16 (4) (April): 561 -573. doi:10.1093/cercor/bhj004.

Gallagher, Helen L., and Christopher D. Frith. 2003. “Functional imaging of ‘theory of mind’.” *Trends in Cognitive Sciences* 7 (2) (February): 77-83. doi:10.1016/S1364-6613(02)00025-6.

Gobbini, Maria Ida, Claudio Gentili, Emiliano Ricciardi, Claudia Bellucci, Pericle Salvini, Cecilia Laschi, Mario Guazzelli, and Pietro Pietrini. 2011. “Distinct Neural Systems Involved in Agency and Animacy Detection.” *Journal of Cognitive Neuroscience* 23 (8) (December 21): 1911-1920. doi:i: 10.1162/jocn.2010.21574</p>.

Greif, Avner. 2006. *Institutions and the path to the modern economy: lessons from medieval trade*. Cambridge University Press.

John B Van Huyck, Raymond C Battalio, and Richard O Beil. 1997. *Tacit coordination games, strategic uncertainty, and coordination failure*. David K. Levine, August. http://ideas.repec.org/p/cla/levarc/1225.html.

John C. Harsanyi, and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Vol. 1. MIT Press Books. The MIT Press. http://ideas.repec.org/b/mtp/titles/0262582384.html.

Knutson, B, C M Adams, G W Fong, and D Hommer. 2001. “Anticipation of increasing monetary reward selectively recruits nucleus accumbens.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 21 (16) (August 15): RC159.

Knutson, Brian, G. Elliott Wimmer, Camelia M. Kuhnen, and Piotr Winkielman. 2008. “Nucleus accumbens activation mediates the influence of reward cues on financial risk taking.” *NeuroReport* 19 (5) (March): 509-513. doi:10.1097/WNR.0b013e3282f85c01.

McCabe, Kevin, Daniel Houser, Lee Ryan, Vernon Smith, and Theodore Trouard. 2001. “A functional imaging study of cooperation in two-person reciprocal exchange.” *Proceedings of the National Academy of Sciences* 98 (20): 11832 -11835. doi:10.1073/pnas.211415698.

McClure, Samuel M, Nathaniel D Daw, and P Read Montague. 2003. “A computational substrate for incentive salience.” *Trends in Neurosciences* 26 (8) (August): 423-428.

North, Douglass Cecil. 1981. *Structure and change in economic history*. Norton.

O'Doherty, John, Peter Dayan, Johannes Schultz, Ralf Deichmann, Karl Friston, and Raymond J. Dolan. 2004. "Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning." *Science* 304 (5669) (April 16): 452 -454. doi:10.1126/science.1094285.

Pagnoni, Giuseppe, Caroline F Zink, P Read Montague, and Gregory S Berns. 2002. "Activity in human ventral striatum locked to errors of reward prediction." *Nature Neuroscience* 5 (2) (February): 97-98. doi:10.1038/nn802.

Rajapakse, JC, F Kruggel, JM Maisog, and DY von Cramon. 1998. "Modeling hemodynamic response for analysis of functional {MRI} time-series." *Human Brain Mapping* 6: 283-300.

Ramnani, Narender, and Adrian M. Owen. 2004. "Anterior prefrontal cortex: insights into function from anatomy and neuroimaging." *Nat Rev Neurosci* 5 (3) (March): 184-194. doi:10.1038/nrn1343.

Rankin F.W., Van Huyck J.B., and Battalio R.C. 2000. "Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games." *Games and Economic Behavior* 32 (2): 315-337.

Rilling, James K, Alan G Sanfey, Jessica A Aronson, Leigh E Nystrom, and Jonathan D Cohen. 2004. "The neural correlates of theory of mind within interpersonal interactions." *NeuroImage* 22 (4) (August): 1694-1703. doi:10.1016/j.neuroimage.2004.04.015.

Rustichini, Aldo, John Dickhaut, Paolo Ghirardato, Kip Smith, and Jose Pardo. "A brain imaging study of the choice procedure." *Games and Economic Behavior* In Press, Corrected Proof. http://dx.doi.org/10.1016/j.geb.2004.08.005.

Schultz, Robert T., David J. Grelotti, Ami Klin, Jamie Kleinman, Christiaan Van der Gaag, René Marois, and Pawel Skudlarski. 2003. "The role of the fusiform face area in social cognition: implications for the pathobiology of autism." *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358 (1430) (February 28): 415 -427. doi:10.1098/rstb.2002.1208.

Skyrms, Brian. 2001. "The Stag Hunt." *Proceedings and Addresses of the American Philosophical Association* 75 (2) (November): 31-41.

Straub, Paul G. 1995. "Risk dominance and coordination failures in static games." *The Quarterly Review of Economics and Finance* 35 (4). The Quarterly Review of Economics and Finance: 339-363.

Tomlin, Damon, M. Amin Kayali, Brooks King-Casas, Cedric Anen, Colin F. Camerer, Steven R. Quartz, and P. Read Montague. 2006. "Agent-Specific Responses in the Cingulate Cortex During Economic Exchanges." *Science* 312 (5776) (May 19): 1047 -1050. doi:10.1126/science.1125596.

Vogeley, K, P Bussfeld, A Newen, S Herrmann, F Happé, P Falkai, W Maier, NJ Shah, GR Fink, and K Zilles. 2001. "Mind Reading: Neural Mechanisms of Theory of Mind and Self-Perspective." *NeuroImage* 14 (1) (July): 170-181.

Völlm, Birgit A., Alexander N.W. Taylor, Paul Richardson, Rhiannon Corcoran, John Stirling, Shane McKie, John F.W. Deakin, and Rebecca Elliott. 2006. "Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task." *NeuroImage* 29 (1) (January): 90-98. doi:10.1016/j.neuroimage.2005.07.022.

Wu, Shih-Wei, Mauricio R. Delgado, and Laurence T. Maloney. 2011. “The Neural Correlates of Subjective Utility of Monetary Outcome and Probability Weight in Economic and in Motor Decision under Risk.” *The Journal of Neuroscience* 31 (24) (June 15): 8822 -8831. doi:10.1523/JNEUROSCI.0540-11.2011.