THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF MATHEMATICS

ANALYZING THE PAYOFFS OF LYING IN VARIOUS SITUATIONS TO DETERMINE IF LYING IS EVER TRULY BENEFICIAL.

ELENA CADENAS
SPRING 2019

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Mathematics
with honors in Mathematics

Reviewed and approved* by the following:

Christopher Byrne
Associate Professor of Mathematics
Thesis Supervisor

Victoria Sadovskaya
Associate Professor of Mathematics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

Robert Trivers, who has pioneered research in deception, has made the statement that lying is never truly beneficial. This paper offers a model to test this claim. The model incorporates signaling into the Prisoner's Dilemma game from classical game theory, thereby allowing players to lie or be truthful about their intended moves, as in the popular T.V. game show, The Golden Ball. Taking into account the psychological and neurological activity of the players while their interaction is occurring, various parameters (confirmation bias, conflict of interest, sensitivity, tells, fear and gullibility) are accounted for to feed into the final credibility assessment of whether the opponent is lying or not. As any game theory model, numerical values are assigned to represent the payoffs the players receive from the various combination of moves that can occur. However, this model utilizes the basic game payoff as well as a collateral biological payoff to account for the stress that an individual's course of action in the game can impose on their body. Analysis considers two cases: both game and biological payoffs in the same units, such as might apply in an evolutionary fitness context, and alternatively, the case when the payoffs are not in the same units, as might apply in an economic context where a trade-off can arise between financial gain and health. For the former context, an overall expected payoff is computed for a variety of scenarios, and for the latter, the Pareto boundary was examined to quantify the trade-off between the payoffs. From the expected payoff calculations, it was determined that lying is never truly beneficial. From the Pareto analysis, conditions were identified that force a trade-off between game payoff and biological payoff. While promising, this conclusion cannot be confirmed until the model is validated. Also, extensions to a repeated game context and other topics for future research are considered.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. Chris Byrne, for always being there for me throughout the entire thesis process. His endless support and willingness to accommodate me into his schedule has made the research process a pleasure. Working with such a charismatic and intelligent advisor has given me a new perspective on research.

I would also like to thank my honors advisor, Dr. Victoria Sadovskaya, for always being willing to meet and give me advice throughout my time at The Pennsylvania State University. Her guidance has given me an amazing undergraduate experience.

Chapter 1

Introduction

There are myriad reasons why people lie, and endless scenarios in which people lie. Individuals lie because they presume it is beneficial to them, or to a loved one; however, is lying actually beneficial when all factors are considered? This paper aims to answer that question by employing a game theory approach to explore various scenarios in which individuals lie [9]. The present scope solely focuses on the classical Prisoner's Dilemma game, although the approach could be used for any game, meaning any multiplayer decision scenario [5].

The Prisoner's Dilemma is a social game where each player can be swayed by personal gain and self-interest of getting the maximum payoff possible. The original scenario considered for this model is that two people are charged with involvement in a crime, are arrested and placed in separate cells (so the two cannot communicate about their thoughts/course of action). To convict one or both of them for the crime, the police have to persuade at least one prisoner to confess. Hence, the options for moves being cooperate with your fellow prisoner and stay silent (C) and defect against your fellow prisoner by turning state's evidence (D). If both confess, then both will receive a modest sentence as part of their plea deals; if exactly one confesses, then the confessing prisoner will be acquitted and receive a reward for aiding the police, and the other will be receive the maximum sentence; lastly, if both cooperate then they both are acquitted due to lack of evidence without a confession. Payoffs are then assigned to each possible outcome. For D, D the payoff is worse than C, C, which is a good payoff for each. For the D, C scenario, the payoff for the player who plays D is the best possible payoff and the payoff for the C player is the worst possible outcome. Hence D is a dominant strategy for both; however, if both players play this dominant strategy then they each receive a lesser payoff than if they both cooperated. Players risk getting the worst payoff (the sucker's payoff) if they cooperate and the other player defects, so mutual cooperation is risky [5].

Table 1: Prisoner's Dilemma Payoff Matrix Example.

		Player 2	
Player 1		Cooperate	Defect
	Cooperate	R, R	S, T
	Defect	T, S	P, P

(T > R > P > S)

Player 1 is the row player and player 2 is the column, so the payoffs are ordered as player 1, player 2.

R = the reward payoff = the payoff associated with mutual cooperation (C, C)

T = the temptation payoff = the maximum payoff possible, D player's payoff in the scenario C, D

P = the punishment payoff = the payoff associated with mutual defection (D, D)

S = the sucker's payoff = the payoff the player who plays C receives in the scenario C, D

Table 2: Numerical Example of Prisoner's Dilemma

		Player 2	
Player 1		Cooperate	Defect
	Cooperate	2,2	0, 3
	Defect	3, 0	1,1

This payoff matrix can be read the same as Table 1.

For this numerical example T = 3, R = 2, P = 1, S = 0.

The Prisoner's Dilemma can be adapted to model many situations and can even be generalized to more than two players. In the present context, the trust required to make cooperation rational is a key feature and the players signal their moves so that the consequences of honesty and deception can be compared. For signaling, the players are not presumed to be separated as in the story for which the game was named, but the assumption of simultaneous moves suffices to preserve the element of risk [5]. If players signal (e.g. announce) their intended moves before each actually makes their move, the important factor becomes who believes whom. If player one defects and player two believes they are going to cooperate, then player two suffers [5]. If player one plans to defect, but signals cooperation and player two detects the deception, then player one suffers [5]. The only way both players win in this game is to both cooperate, both getting a net positive payoff [5]. Telling the truth in this game would mean telling the other player what your move is going to be and actually playing that move. Lying would mean telling the opponent that you are going to either cooperate/defect and then doing the opposite. There could be reasoning for each scenario as seen in the TV game show Golden Balls, which models the Prisoner's Dilemma [26].

In this game show, the players are each given a choice of two balls. One contains the word "split" while the other one says "steal" inside. If both players pick split, then they split the prize money in half. If one chooses to steal and the other split, then the player who chose steal receives the full sum of prize money. If both players play steal, then they both leave with nothing. The players are allowed to talk strategy before making their moves. Take into consideration player 1 (P1) and player 2 (P2). If P1 promises to split with P2 then P2 can lie and play steal since they think P1 is going to play split to get the largest payoff possible. However, in this case both players could be lying and play steal leading to no reward for either player. P1 can claim they are going to steal no matter and then split the reward with P2 later. This way P2 is left to only play split in hopes to receive money after the game. P1 could have been lying and end up playing split since they were planning to split the reward with P2 all along [26].

Building on the multi-dimensional player models, specifically the signaling and belief updating of Byrne and Kurland, this thesis exploits concepts from psychology and neuroscience to model the mental processes of each player leading up to their moves [3].

Chapter 2

Background and Related Research

Lie Detection

Lie detection has been utilized in legislation to ensure that the truth, and nothing but the truth, is used in various cases [7]. Many polygraphs are not connected to the head of the individual being tested, but rather monitor the physical autonomic nervous system responses such as heart rate and blood pressure, which mostly monitors the anxiety level associated with lying versus the actual detection of brain stimulation due to lying [7]. To better understand the origin of lying patterns in the brain, the brain ought to be monitored to determine the cortexes that are active during deception. This will allow for a more accurate detection of deception since the individual will not be able to consciously control their anxiety response associated with lying. The University of Pennsylvania performed a study to determine the active cortexes for deception [27]. A functional magnetic resonance imaging machine (fMRI) was utilized to determine that the active cortexes during deception are the anterior cingulated cortex (monitoring errors), the dorsal lateral prefrontal cortex (controlling behavior) and the parietal lobe (processing sensory input). This machine monitors the blood flow in the brain to determine where the most activity is occurring by magnetizing molecules in the brain to make them resonate and be detected.

Images are captured by a computer that then uses mathematical formulas to process the information [27].

Using fMRI could be a more accurate way to detect deception instead of the typical polygraph; however, it costs more and is more difficult to test subjects since they must lie still while being imaged. In light of this new methodology, some individuals are looking to pay to prove to their spouses that their loyalty is true. Thinking of this in terms of payoffs, an individual would have to determine if they think the $10,000 is worth demonstrating their loyalty if, in the end, their spouse may still not believe their word. Hence, there is no guaranteeing that this testing will make their spouse not doubt their loyalty. Additionally, if their spouse does not trust them in the first place then there is a deeper issue occurring that should be focused on instead of deception [27].

Consciousness and Unconscious Deception

This study aims to analyze the decision to lie by constructing a model that includes both the objective and perceived consequences of lying, especially how these decisions and consequences are processed in the brain. Lying can be a conscious or unconscious act [18, 23, 29]. Research by Marvin Minsky and Robert Trivers has consistently demonstrated that when consciously lying to yourself, or others, the left side of your brain is mainly at work [18, 29]. The right brain controls the unconsciousness. Deciphering between conscious and unconscious lying is crucial. There is a time delay between conscious and unconscious, so it is hard to sensor the world. This is also where déjà vu comes from and why memories can seem so real. Since the neural processing of actual time is delayed, memories can seem like they are actually occurring [18, 29].

To verify the findings of the University of Pennsylvania's study, researchers at the University of Baylor conducted an experiment on bargaining with sellers and buyers [27, 14]. The sellers were categorized as either strategists (actively lie to profit), incrementalists (do what is fair), or conservatives (would rather make a sale than profit, so they ask for under value) [14]. An fMRI machine was also utilized in this study to discover the active parts of the brain for each type of seller. The dorsal anterior cingulate cortex, retrosplenial cortex and rostral prefrontal cortex were the main cortexes that lit up in the brain of strategists lying. For the incrementalists, the dorsal anterior cingulate cortex was found to be lit up when they lied. The conservatives had a lot of variation in the fMRI scan results, so no information could help formulate a conclusion on the active parts of their brain when lying. Using these fMRI scans, it was also discovered that there is directional connectivity between the brain cortexes. That means that these cortexes of the brain are not mutually exclusive since they all can communicate with each other [14].

The findings on brain activity in strategists and liars align with the results of the study done at the University of Pennsylvania [27]. They both found the anterior cingulate cortex lit up upon lying, which is part of the limbic system [14, 27]. The retrosplenial cortex is the odd cortex out of the University of Pennsylvania research; however, it is a part of the posterior cingulate cortex dealing with memory, so it makes sense that it was active as well [14, 27]. For the purpose of this research, the retrosplenial cortex and the anterior cingulate cortex will be grouped to just be the cingulate cortex [30]. The rostral prefrontal cortex found in this paper corresponds to the prefrontal cortex indicated in the University of Pennsylvania study, which controls an individual's behavior as well as monitors the behavior of others [14, 27].

Since lying can induce stress, pleasure and pain, one would assume there would be a correlation between the cortexes involved in these physiological responses and the cortexes involved in lying [17].

The hypothalamic-pituitary-adrenal axis, the amygdala and the prefrontal cortex are the parts of the brain associated with these physiological responses according to multiple studies done at NIH [30]. The prefrontal cortex is listed in all of the studies considered thus far, which has the function of voluntary muscle control, higher order thinking, memory, and speech production [17]. The anterior cingulate cortex is part of the limbic system, which also contains the amygdala and hypothalamic-pituitary-adrenal axis [22]. These all deal with emotional and cognitive processing such as your pleasure center. The parietal lobe deals with the senses and receiving sensory information, so it makes sense that this is involved in lying, but not in the physiological responses induced by stress, pleasure and pain [17].

Self-Deception

There is a motive behind every lie. People mainly use deception for the benefit of themselves or others. Individuals lie when they think they are doing something “good” even when they are not [29]. There can be a conflict of interest when lying, since something can benefit one person the individual cares about, but harm another at the same time [4]. Thus, the projected benefits are always analyzed prior to deciding to lie [4]. An individual might believe that a lie is beneficial, when the lie actually comes with more consequences than benefits. Hence, the actual benefits and consequences of lying can be quite different than the projected results [8].

In the case of self-deception, the motives are normally to confirm a bias or help project a positive self-image [28]. Whether the projected outcome be that the individual receives more friends out of the lie, or that they want to feel good about themselves; the projected outcome could match or be extremely different than the actual outcome. Individuals could lie to themselves, believing that they can lift a lot of weights since they believe they are so strong and then end up injuring themselves because they were overconfident that they could lift more than they could. Overconfidence is one of the most common forms of self-deception [29]. As in this example, many times it is people thinking they can do something even though they cannot. There is also functional overconfidence where you believe you can do something before trying it [28]. This is evolutionary adaptive because it causes people to try more things and build up skills [28]. Taking functional overconfidence into account, you do not want to bite off more than you can chew, but on the other hand it is not always destructive meaning sometimes it can get you to do something you might not have previously tried.

When people believe in success, but experience contradicts them, they lose faith. Thus, the problem with overconfidence. People should believe in smaller goals so that they succeed. Then, their capacity of faith in themselves gets larger and they can then achieve those larger goals that they would have originally failed if attempting it from the get-go [29].

Another example of projecting a positive self-view through self-deception is described by Leonard Pitts, Jr. [21]. He states that if an individual pays a lot for a forgery then they are not very likely to admit it is a forgery [21]. That is, there is a positive correlation in the fact that the more money paid for the forgery, the more likely the individual is to lie that it is not a forgery [21].

An additional motive for self-deception is that the individual is more successful at lying to others if they too believe their own lie [28]. This has become adapted since it is evolutionary beneficial [28]. As displayed in Voltaire's Candide, people lie to themselves for benefit of pretending something is not as bad as it actually is [31]. In other words, they lie to themselves to portray a false reality as a mode of comfort. In this scenario, the benefits outweigh the costs of not knowing what the truth is [31]. Once you get caught up in a lie it is hard to convince yourself that it is actually a lie and not reality. This type of self-deception is denial [29]. Another form of denial is denying that you care about someone because you are afraid of getting hurt. Many people lie to themselves in relationships because they think it is easier to not feel any emotions than to potentially get hurt. The issue with this is that they could be missing out on something amazing and miss feeling unexpected emotions.

A third form of self-deception is confirmation bias [28]. Confirmation bias is the tendency to interpret new evidence as confirmation of one's existing beliefs or theories [29]. Many individuals will ignore information that goes against their existing beliefs and lie to themselves disregarding that information as valid. If something goes against their preexisting belief, then they also will lie to themselves thinking that that information is not true. This just solidifies their existing bias and will strengthen it making them less open minded to anything that goes against this bias [29].

In accordance with the forms of deception listed above, Trivers states that there are nine categories of self-deception [29]. The first is self-inflation, where people put themselves in the top half of all positive distributions and the lower half of all negative attributes. They do this by suppressing neural activity in the region where the activity takes place. The second is derogation of others. This is where an individual derogates an entire group instead of changing their self-image. People have adopted this mechanism as a defensive strategy when they feel threatened. The third is in-group/out-group associations. This means that once the individual defines an out group, they degrade them and put more positive traits on members of the in-group [29].

The fourth category is biases of power. With the feeling of power, people are less likely to take others' viewpoints and are more likely to center thinking on themselves. Moral superiority is the fifth category Trivers proposes. When determining our value to others is when this category is most prevalent. The sixth is illusion of control. We believe we have greater ability to affect outcomes when deceiving falls in this category. The individual responds to a lack of control by generating false data that would give them greater control. The seventh category is construction of biased social theory. Theories regarding our immediate social reality fall into this category. An example is a husband and wife, with one being an altruist and the other being selfish and them both disagreeing on who is who because they have different social realities. The eighth category is titled false personal narratives. This is where an individual enhances their view of themselves and derogates others. Lastly, the ninth category is unconscious modules devoted to deception. Two or more activities can occur with no interaction, making them independent conscious/unconscious modules. An example of this is stealing things from people unconsciously while interacting with them because they are getting something from the interaction, so the individual wants to get something in return [29].

Deceiving Others

The main two bases of motivation for lying to someone else are to benefit yourself or another person (not necessarily the person you are deceiving) [15]. The projected benefits of the lie are dependent on what the lie is [8]. For example, if you are lying to someone because there is only one donut left and you want to eat it then the projected benefit would be to eat the donut. The lie could be telling the other person that the donut is stale, so that they do not want to eat the donut and leave it there. The risk of this lie is that if the person finds out you were lying then they could eat the donut in retaliation. They also could just think lesser of you and believe you are selfish wanting the whole donut for yourself when you both could have split the donut. This displays how when lying, people tend to focus on the immediate reward versus the high long-term cost (instant gratification of eating the donut versus having your integrity jeopardized) [8].

The individual might not take into account the risks and long-term costs that could potentially occur from lying [8]. When about to use deception, all of the potential benefits and risks should be considered. A major negative impact that is associated with deception is stress on the immune system due to feeling guilty and shameful [29]. Alongside making the original projection, when reality does not live up to the forecast, the individual should correct their forecasting process for future projections; however, temptation comes in to play for the individual to pretend that everything is working fine so that the individual continues to use deception even though the associated negative impacts might outweigh the positive impacts [8].

The positive impacts are what most individuals focus on when lying demonstrating why lying evolved [15]. Minsky's view is that lying was created by human learning based on reward and reinforcement where Trivers claims it was created because humans seek pleasure [18, 29]. From Minsky's description of agencies, we see that pleasant feelings of accomplishment and disagreeable sensations of defeat are the main factors of how our higher-level agencies make summaries [18]. Thus, Minsky and Trivers maintain similar views of humans being pleasure seeking creatures [18, 29].

With the fact that people lie for some pleasant outcome, another source to consider is genetics. It has been found that women have been found to be more compassionate towards those who have lied to them. This could be because of genetics, or the way they were raised to be (nature vs. nurture). The genetic makeup of an individual could impact how lying plays out for them. If an individual has a weak immune system based on their genetic makeup, then lying is more of a risk for them than someone with a strong immune system. The immune system is also impacted by how you are raised due to it being shaped from outside resources that are introduced to the body. This shows that nature and nurture both have compelling arguments to how an individual is created and evolves to having tendencies of deception [29].

When deceiving other people, I will define the following two forms of deception: active and passive lying. Active lying is altering the truth, while passive lying is omitting the truth. Passive lying is very common when keeping secrets from people. The donut example is a way where the truth was altered, thus it is active lying. These are the two types that all deception of others stems from. Also, multiple scenarios can contain active and passive lying. For example, when lying to your doctor about your current condition, you could omit symptoms (passive lying) or give a false representation of the symptoms you have (active lying). Having this classification system makes it simpler to relate various scenarios of deception.

Trivers claims that if everyone deceives, then it is never truly beneficial because we are just living in a circle of lies [29]. In the movie “Life is Beautiful,” the man lies to his son to keep him alive in the concentration camps [1]. He tells him that it is a big game of hide and seek so that the son never is found [1]. This challenges Trivers’ argument that lying is never beneficial because in this scenario one would think that both the son and the father would claim that this lie benefited them [29]. Another contradiction of Trivers’ argument is from Leonard Pitts, Jr. He argues that it is beneficial to lie to those who have lost loved ones in war, since it is better to think that your loved one died serving their country rather than dying because of an oil dispute [21].

Deception of others is prevalent in nature as well. An example is how some species of nonpoisonous butterflies mimic poisonous ones so that they are not eaten by predators [29]. The more frequent these deceivers act, the more they begin to diversify themselves to better avoid detection. Hence, the deceivers evolve quickly to maintain their success. Another form of deception in insects is caterpillars’ emission of pheromones (a chemical produced by an animal that is projected into the environment that affects the physiology of other species) to make them smell like ants and act as ants so that when a predator attacks, then the queen ant will protect them over the actual ants [29].

The Positive and Negative Aspects of the Results of Lying

When analyzing the benefits and consequences of lying, all individuals effected by the lie must be considered [8]. In consideration of a net payoff for the scenario as a whole, there will be weights on each individual based on their connection to the individual who lied [29]. The more connected to the liar, the higher the weight. Due to the chain reaction of the impacts of lying, there will be individuals affected by the lie who are not directly connected to the liar but are connected to a person who is connected directly to the liar [29]. These indirect contacts will have a much smaller weight on the overall end payoff of the scenario, since they are farther removed from the original lie itself. From the Folly of Fools, Hamilton's rule states that the benefit of an altruistic act toward a relative multiplied by the relevant degree of relatedness must be greater than the cost suffered by the altruist in order for natural selection to favor altruism [29].

A significant factor in the net payoff is the immune response of the individuals involved in the lie [29]. Lying can cause a lot of stress on the individual lying because they can feel guilty or shameful [29]. This stress causes the immune system to focus on dealing with the stress the body is going through, which makes it easier for the body to be attacked by a foreign antigen [2]. An antigen is a substance introduced to the body that causes the immune response to react by creating antibodies (the body's cells that protect it from infections etc.). The depressed immune function lowers brain activity since it takes a lot of energy to lie, hindering the allocation of energy to the brain so a successful lie is less probable [2]. Repression of the truth from the conscious mind also lowers the immune function because the brain activity takes a lot of the body's energy supply [29]. Once someone becomes a pathological liar, lying no longer requires a lot of energy because it is almost a natural response to a stimulus [20].

People become accustomed to deceiving themselves and others. Thus, deception can become a subconscious act [20]. As described by Marvin Minsky, an example of this is giving someone something so you think that you need to take something from them as well, and you subconsciously take something [18]. The more one lies, the less active their amygdala becomes each time a lie occurs, so they are no longer stressed or having a fear of the potential outcome of the lie [20]. When the negative physiological factors of lying are removed, deception becomes addictive [20]. Along with the removal of these factors, the most salient things tend to happen, so neural pathways fire quickly when used more [20]. Hence repetition of this firing makes it more likely to fire again.

When someone has become accustomed to lying, their lies are less likely to be detected because they will have less of an emotional connection to lying [20]. People can detect lies through the liar's nervousness, control (of their words and actions), and cognitive load (the brain's effort in processing specific information) [19]. These cues make it easier for others to become aware that a lie is taking place. If the liar picks up that the other person realizes they are lying, then the probability of detection also can impact if the liar will switch from lying to telling the truth.

Theory and Research Specific to Model Structure

Belief, Desire, Intent Philosophy

The belief, desire, intent philosophy (BDI) turned out to be crucial in creating the base for this model instead of being a parameter since it describes the mental process that goes into an action [32]. Each player's belief is what they believe the other player is going to do [32]. Each player's desire is the payoff they want (their aspiration level in economic terms), which in the case of the Prisoner's Dilemma desire can be reward, temptation, or punishment (or the sucker's payoff, but normally, nobody wants that) [5, 32]. Reward is the desire for a fair share of a mutually cooperative (C, C) outcome, and so could lead a player to play C, but could also lead to a D move if a player assesses (correctly or not) that their co-player is going to play D [5]. Temptation is the temptation to have the maximum payoff possible, and a player must play D to get it, and hope that their co-player plays C [5]. Punishment, a relatively low desire, can also lead a player to play D, because it guarantees the punishment payoff without having to worry about the other player's move [5]. As a starting point, to simplify the number of possible outcomes of play, the desire category has been set at reward for all possible combinations of play. Each player's intent is the action they intend to play (C or D) and whether they intend to send a truthful or deceptive signal [5, 32]. Thus, this model has 2-dimensional intent, compared to early BDI models with 1-dimensional intent [5, 32]. The signal is what a player signals to the other player that they are going to play [32].

This game is noncooperative insofar as lacking binding agreements and dropping the cooperative assumption of collective rationality, but, like cooperative games, some level of communication is assumed and is modeled by each player signaling a move and signals can be truthful or not [9]. Re-allocation of payoffs after the game is not actually done, also adhering to traditional non-cooperative game theory assumptions, but reallocation should be possible to make promises of side payments plausible, as in the Golden Ball [6, 26]. For this model we have set the desire of both individuals as reward to eliminate hundreds of possible scenarios (which can be further analyzed in later work). With the desire of both individuals strictly being reward and since they lack binding agreements and therefore have to assess each other's truthfulness and intended move, the best strategy for both players is *anticipatory tit for tat*; i.e. for each to mimic the move that they forecast the other player will play [3].

Signal Detection Theory

The next relevant theory is the signal detection theory. This theory is how individuals make decisions that involve perceptual and behavioral risk, which deals with the connection of the senses and how the brain processes them then the individual finalizes a decision [16]. The brain lobes for this theory are also the parietal lobe and prefrontal cortex since it deals with sensory and higher order processing [17].

Extensive research has been done to create a mathematical model of this theory that separates the perceivers behavior into sensitivity and bias. There are four parameters included in the model as follows: payoff parameter, base rate parameter, similarity parameter and the sensitivity parameter. The payoff parameter is based on correct detection, missed detection (aka false rest), false alarms (incorrect detection), and correct rejections. The base rate parameter is the perceivers probability of encountering targets versus foils, in our case people telling the truth and people lying. The similarity parameter is what targets and foils look like (truthful people or liars). External factors can have an impact on this parameter as well. The last parameter is the sensitivity parameter, which affects the players biases. the factors of the sensitivity parameter are the liberal and conservative criterion. The liberal criterion decreases missed detections but increases exposure to false alarms. The conservative criterion decreases false alarms but increases exposure to missed detections. Thus, the perceiver needs to optimize their criterion location to maximize the expected outcome [16].

It has been shown that there are multiple combinations of bias and sensitivity that lead to the same accuracy; however, the maximum accuracy has been found to be at a level of zero bias [16]. The formula for bias that they created with β=bias, c= bias criterion measure and d=sensitivity measure is: $\beta = e^{(cd')}$ [16]. From this function, the optimal bias equation was derived.

$\beta_{\text{optimal}} = (1 - \alpha)/(\alpha) \bullet (j - a)/(h - m)$ with α=base rate, j, a, h, and m being payoffs for correct rejections, false alarms, correct detections and missed detections respectively [16]. The derivation of this equation can be found in the paper written by Lynn and Barrett [16]. In order to incorporate confirmation bias along with the other psychological factors considered, this formula will be utilized as a guide for the equations created.

Projection

Psychological theories were discovered and if found to be applicable to decision making and deception, the brain cortexes utilized in these theories were researched. The first theory that applies to decision making is the projection factor. The projection factor is when an individual projects his/her feelings and beliefs on another individual [10]. This factor is more likely to occur if the individual has negative feelings or emotions [10]. It could also be more likely with positive emotions/feelings as well if they are very strong and the person is highly biased [10]. The parts of the brain involved in projection are the parietal, temporal, occipital lobes, and prefrontal cortex [24]. Based on the parts of the brain involved in deception, these lobes correspond quite well with the lobes already considered in our model. The temporal and occipital lobe are both lobes that deal with sensory input, hearing and seeing respectively [17]. These lobes make up the portion of the brain called the cerebral cortex that deals with all sensory processing [17]. For our model, since we are focusing on the overarching processing of the senses, we will take into consideration the parietal lobe and prefrontal cortex and omit the temporal and occipital lobes. Due to the ventral and dorsal pathways connect all the lobes and the parietal lobe and prefrontal cortex connect all of the information received from the occipital and temporal lobes [17].

Confirmation Bias (CB_j)

Confirmation bias (CB) is the tendency to reaffirm one's prior belief(s) [12]. In the present model, players' prior beliefs about each other are based on projection [10]. As discussed in recommendations for future work, prior beliefs could reflect history in repeated game context. CB is a function of initial belief and initial deception/honesty [12]. This also includes the projection factor where the player is projecting their beliefs onto the other player, which could be projecting in two different ways [10]. First off, they can be projecting their intent (to play C or D) to be their belief of what move (C or D) the other player is going to play [10]. Additionally, they can be projecting their deception/honesty onto the other player, believing the co-player to be a liar if they are a liar, and believing the co-player to be signaling honestly if they are signaling honestly [10]. With CB, activity can be seen in the prefrontal cortex because of it being higher order processing [24]. The scale categorization for this parameter is 0 being a strictly conservative bias and 1 being a strictly liberal bias. A conservative bias meaning that the individual is not strongly fixated on their bias [16]. This decreases false alarms but increases exposure to missed detections [16]. For the liberal bias, the individual is highly opinionated and attached to their bias decreasing the amount of missed detections, but also increasing their exposure to false alarms [16].

Conflict of Interest (COI_j)

COI considers the tendency to believe what you want to believe (literally, COI skews judgement to favor preferred outcomes) [4]. Therefore, its effect would be most extreme when desire is for a high temptation payoff but can even result in the reward from mutual cooperation being perceived as more likely than it truly is [4]. Individuals are normally not aware when their judgement is being skewed, so in these situations this falls into their unconscious brain [4]. The prefrontal cortex's activity is monitored here since the player is debating their optimal strategy [17]. For the probability of the player to be completely blinded by the temptation payoff the weighted value is 1, and when it is 0 the player retains perfect judgement in the face of the temptation payoff regardless of the possibility of desirable outcomes.

COI is based on how "starry eyed" the individual is. When the opponent (player not j, "~j") signals C, this give hope to player j of getting the temptation payoff, which as mentioned before is the highest possible payoff. Once player ~j signals C, the final COI is decided based on player j's intent. COI is either .6 if player j intends to play C and .9 if they intend to play D (since they are very starry eyed and really want player ~j to be telling the truth and actually play C). When player ~j signals D, player j still has a slight COI (.2) since player ~j could be lying and actually play C.

Tells (T_j)

T represents the tells that players give off that alert the other player whether or not they are lying. Furthermore, the tells of player j, T_j, are the tells that they are receiving from player ~j. These tells are like the tells that poker players have [19]. For example, the nervous energy generated from lying leads the individual to release that energy in one way or another such as moving their tie, or tapping their finger, or adjusting their glasses [19]. Another common example of a tell is with police questioning individuals. The individuals give off tells that they are lying whether it be excessive swallowing, having extra delays between questions, or performing nervous tendencies such as fixing their hair [19]. If player j is lying, then the tells they give off contribute to player ~j thinking they are lying. T is associated with activity in the parietal lobe due to tells being sensory output information that the player gives off and the input processing of this sensory information [17]. With this parameter, 0 means no indication of deception at all, and 1 is the strongest possible signal of deception.

Tells is dependent on the player's sensitivity and they have an inverse relationship. If the player is lying, then they are not very sensitive and giving off a lot of tells (.8) since they are nervous about lying. On the contrary, if the player is telling the truth then they are highly sensitive correlating to not a lot of tells (.2).

Sensitivity (S_j)

S is a player's sensitivity to Tells, discussed above as tell-tale signs of honesty and deception, so this parameter is the probability that the individual will pick up on those signs, which is dependent on their sensitivity to these signs. Due to the fact that if player j is lying, they will be focused on ensuring that their behavior is not giving away that they are lying, S is inversely proportional to an individual's own deception (furthermore, the tells they give off) [28]. If someone is telling the truth then they are more likely to be more aware of their surroundings, allowing them to be more sensitive to the other player's signals and body language [19]. S deals with the parietal and prefrontal cortex as well as the limbic system [17]. This is due to the fact that the individual has sensory input (parietal lobe), which is analyzed by the prefrontal cortex on their interpretation of the information [17]. The limbic system dealing with stress impacts this parameter because if the individual is more stressed then they are less likely to pick up on the tells from the other player [22]. For this scale, 0 means player j has no sensitivity; i.e. a 0% probability of picking up on player ~j's Tells. 1 is player j having a 100%

probability of picking up on player not j's signals of honesty or deception, a probability of utmost sensitivity.

Fear (F_j)

F is a function of the sucker payoff. Opposite of COI, F is the fear of getting burned by playing cooperate and getting the sucker payoff due to the fact that the other player plays defect. This leaves the player that cooperated with nothing, and the player who defected with the maximum payoff amount. The limbic system and prefrontal cortex are active with F due to the stress the player is going through and them having to analyze the situation to determine if playing cooperate is worth the risk of getting the sucker payoff [17]. Here, 0 represents the lack of fear that the individual is going to get the sucker payoff and 1 is the maximum level of fear the player could have about getting the sucker payoff [13]. Since fear affects risk aversion, it acts as the threshold of the lie detector. This value was compared to the average value of a player (avg), and if the avg was greater than or equal to the F value then the player believes the opponent is telling the truth, and if the value is equal to or below this threshold value then the player believes the opponent is lying

Gullibility (G_j)

G is the gullibility and tolerance of the individual. This corresponds to the likelihood of detection because the more gullible the person is then the less likely they are to believe the other player is lying and the more likely they are to believe whatever the opponent signals to them. The limbic system and prefrontal cortex are activated with this parameter because if the player is utterly suspicious of the opponent, then their stress levels in the limbic system will be high and gullibility and suspicion are highlighted in the prefrontal cortex [17, 22]. The representation of strength for G is 0 being completely is lying regardless of all other factors) and 1 being completely r is telling the truth regardless of all other factor). The default probability of gullibility is .5 where an individual has an equal probability of believing or not believing the other player's signal. This set value is the mean of sensitivity, since the sensitivity of the player can either increase or decrease from the gullibility of the person. This is due to the fact that a person's level of suspicion/gullibility will have them entering the game play with an already active brain. When accounting for this, the endogenous variables are calculated starting with this mean as the base value.

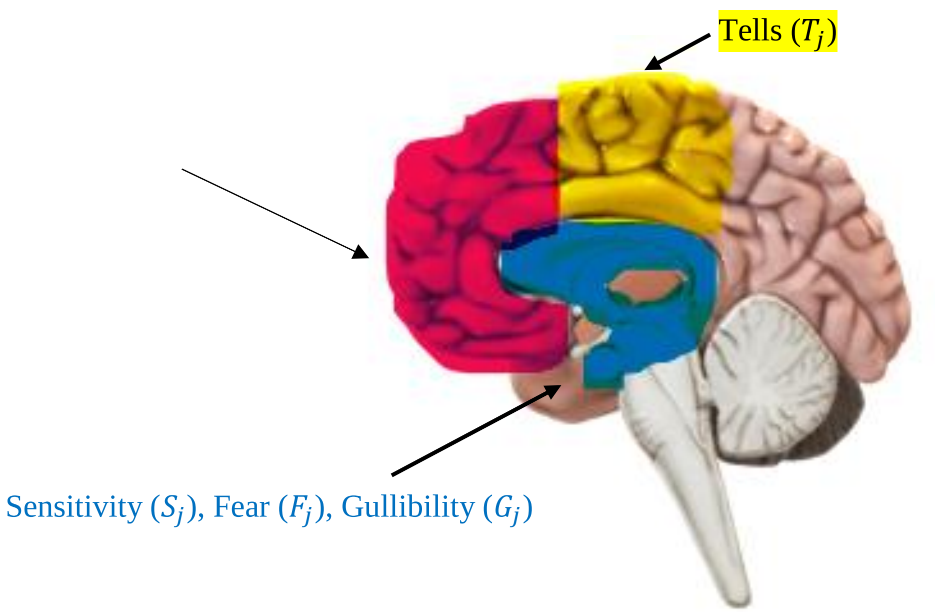

Figure 1: Brain Model

This model of the brain shows its structure where the parts of the brain involved in lying have been highlighted. The pink area is the frontal cortex, containing the prefrontal cortex; the yellow area is the parietal lobe; and the blue area highlights the components of the limbic system. The corresponding parameters for each brain area are labeled as well.

Payoffs

In the world of game theory, payoffs are preferences over outcomes, and therefore also incentives for an individual to make a decision [13]. Player j's *best response* to player ~j's move is the move that maximizes player j's payoff [13]. In this model, we are no longer solely focused on the direct payoffs from the game, but also the biological biproducts of the player's move and decision-making process (i.e. what parts of the brain are calling the shots). We will consider payoffs for various parts of the brain where each part of the brain is a decision-making algorithm that can be reinforced or discouraged (the prefrontal cortex, limbic system, and parietal lobe). This compilation of regional activity payoffs in the brain comprise the collateral biological payoff (based on the collateral damage lying does to your body).

Along with the biological payoff, the traditional game payoff (the typical physical game theory payoff) must be accounted for [6]. The game payoff is established based on the final moves of the players and is in the units of preference in outcome [6]. The biological payoff is based on how the play of the game was processed in the brain, specifically in the prefrontal cortex, parietal lobe, and limbic system (the areas of the brain active in lying scenarios). The activity in each brain location ranges from -50 to 0 with -50 being activity causing stress on player j's body and 0 meaning the activity in this region has no impact on player j's physiological stress.

Chapter 3

Model Specifications

Model Overview

The overall structure of the model mimics the Golden Ball TV game show: two players engage in a Prisoner's Dilemma game, but before they actually make their (simultaneous) moves, they first signal to each other (C or D) [26]. Each player's signal can be truthful (i.e. their true intended move), or false. Within this overarching framework, numerous detailed variables are included to model the psychological and neurological concepts discussed in the preceding section. Moreover, these variables are updated during the signaling phase, paralleling Byrne and Kurland, to model the mental dynamics of the player's assessments of each other and possible changes in intent, culminating the player's moves [3]. The variables are defined below followed by the formulas for endogenous variables, and then there are specifications of payoffs. The model specification includes two annotated examples to illustrate the variable updates and payoffs that result from different initial conditions and concludes with an abbreviated quick reference guide listing all variables.

Exogenous Variables

Gullibility (G)

Gullibility is the mean for the player's overall sensitivity. G_j = .5

Fear (F)

Fear was utilized as the determination threshold of the individual believing the other player was telling the truth or not. Hence, F was held constant for all players. F_j = .5

Tells

T_j = 1 - $S_{\sim j}$

If player ~j is lying, then $T_j = .8$ (player ~j is giving off a lot of tells).

If player ~j is telling the truth, then $T_j = .2$ (player ~j is giving off minimal tells).

Endogenous Variables

The formulas for the endogenous variables require lengthier explanation then the exogenous input variables. For clarity, the formulas are stated first, and they will be explained in the paragraphs that follow.

Final Sensitivity(S)

Equation 1: Player j's final sensitivity

$$Sf_j = (S^2_{\sim j} + (1 - S^2_{\sim j})(.5 + .5(Truth_{\sim j}))) * (S_j)$$

Final Confirmation Bias (CB)

Equation 2: Player j's final confirmation bias

$$CBf_j = .5 + .5(Truth_j)(CB_j)$$

$Truth_j$=1 if player j is telling the truth

$Truth_j$=-1 if player j is lying

Final Conflict of Interest (COI)

Equation 3: Player j's final conflict of interest

$$COIf_j = .5 + .5(COI_j)(Sig_{\sim j})$$

If $Sig_{\sim j} = C$ then set $Sig_{\sim j} = 1$

If $Sig_{\sim j} = D$ then set $Sig_{\sim j} = -1$

If $Sig_{\sim j} = C$ and $I_j = D$ then COI_j = .9 (player j has high hopes of temptation payoff).

If $Sig_{\sim j} = C$ and $I_j = C$ then COI_j = .6 (player j has some hope of temptation payoff).

If $Sig_{\sim j} = D$ regardless of I_j, COI_j = .2 (player j has little hope of temptation payoff).

Final Sensitivity Equation (1)

Final Sensitivity (Sf) was calculated using Equation 1, which is the weighted average (i.e. convex combination) of total belief estimate (= 1 for full credibility) and tells estimate (starts at .5, since 50% chance player ~j is telling the truth). The player's response to tells' weight on total belief is cool factor = 1 (so sensitivity very close to one will persuade them to believe the truth.

"$cool\ factor_j$" =S_j^2 (high sensitivity correlates with a very cool character)

$credibility_{\sim j}$ = degree to which player j believes player ~j

$tells_{\sim j} = truth_{\sim j} * nervousness_{\sim j}$

$nervousness_{\sim j} = 1 - sensitivity_{\sim j}$

$response_j\ \ to\ T_{\sim j} = T_{\sim j} * S_j = truth_{\sim j} * nervousness_{\sim j} * S_j$

Final CB Equation (2)

Equation 2 was developed based on the second part of the convex combination of Equation 1. Starting with the gullibility value of .5, the confirmation bias was added. The truth of player j was multiplied by their confirmation bias then by .5, so that the values range from 0 to 1. Since 1 is the numerical representation of truth and -1 is the value of lying, the starting value of .5 can either increase or decrease. These factors were the only ones included since an individual's bias is solely based on their actions.

Final COI Equation (3)

The same format for Equation 2 was used for Equation 3. However, the factors multiplied by .5 were switched to be player j's COI and the signal of opponent ~j. The signal of player ~j is 1 when the signal is C and -1 when the signal is D. This allows for the fluctuation of the final value being between 0 and 1.

$Sig_{\sim j} = C = 1$ if player ~j signals C

$Sig_{\sim j} = C = -1$ if player ~j signals D (hence negating from the base .5 making player j more likely to believe player ~j is lying and will actually cooperate)

Signals and Cognition

Initial Desire

The initial desire is held constant as being the reward payoff (R) for each scenario.

Initial Intent

The initial intent varies between C and D for each individual, which was assigned per scenario based on the various combinations of belief, intent, and signal for each player.

Initial Belief

The initial belief (Bi_j) of the player is dictated by projection of that player's initial intent (I_j). $Bi_j = I_j$

Signals

Each player signals (Sig) C or D depending on their initial intent and truthfulness.

If truthful: $Sig_j = I_j$

If deceitful: $Sig_j \neq I_j$

(For $Sig_j = C\ or\ D$ and $I_j = C\ or\ D$)

Lie Detection

Fear was compared to the average value for player j.

$$avg_j = \frac{Sf_j + CBf_j + COIf_j}{3}$$

The credibility assessment (CA), which is if player j believes player ~j is telling the truth (TR) or lying (L), was determined based on the following criteria.

Equation 4: Player j's credibility assessment

$$CA_j(I_j, h_j, I_{\sim j}, h_{\sim j})$$

If $avg_j \geq F_j$ then $CA_j = TR$

If $avg_j < F_j$ then $CA_j = L$

Updated Beliefs

The final belief (Bf) for each player is determined based on CA.

If $CA_j = TR$ then $Bf_j = Bi_j$

If $CA_j = L$ then $Bf_j \neq Bi_j$

Updated Intent (moves)

Updated intent becomes the actual move of each individual, so the variable represents both and is determined based on Bf.

$Bf_j = move_j$

Payoffs

$\pi = game\ payoff$

$bio\ \pi = collateral\ biological\ payoff = PF + LS + PA$

(PF = payoff associated with the prefrontal cortex, LS = payoff associated with the limbic system, PA = payoff associated with the parietal lobe)

$net\ \pi = \pi + bio\ \pi$

Explanation of Data for Scenario 2

The data for scenarios 2-17 is displayed in the tables in the following section, but this is an example of how the data can be interpreted for each scenario. Starting with initial belief and ending with net payoff, the following is an explanation of all the different variables for player 1 (P1) and player 2 (P2) in scenario 2. Describing P1 first, they believe that P2 is going to cooperate (C) along with signaling C to P2. They are telling the truth because they intend to cooperate. Their COI is .6 since they plan to C, but P2 signals C, so there is a slight hope of getting the temptation payoff. CB is fixed at .8 meaning that they have a strong confirmation bias that P2 is going to be acting in a similar manner to them. P1's tells are low (.2) since P2 is telling the truth and not sending off many tells. Since they are also telling the truth, their sensitivity is high at .8. Gullibility and Fear are both held constant at .5 meaning P1 is in the middle of being gullible versus suspicious as well as having an average fear of getting the sucker's payoff. Equation 1 was utilized to determine the final S of P1 to be .6, the final CB to be .9 and the final COI to be .8. These three values were then averaged to get .8, which was compared to F (.5) to determine if P1 thinks P2 is telling the truth/lying (P1's credibility assessment/CA). Since avg1 > F, P1 thinks P2 is telling the truth, which is a correct (R) assessment (A). Thus, P1 believes P2 is going to play C, so P1 plays C. Due to having a COI of .6, the prefrontal lobe payoff (PF) of P1 is -30, since it is putting stress that is a little above average on P1's body. The stress placed coming from the limbic system (LS) is 0 since P1 is being truthful. Since tells are low, the sensory information entering the parietal lobe (PA) is low yielding a -10 payoff. These three payoffs were added to get the total biological payoff of -40. The game payoff for P1 is 100 since the final play of moves was C, C meaning both players split the reward. The net payoff for this individual (the game plus the biological payoff) is 60, but this net value is only accumulated if the game and biological payoffs are in the same units.

For this scenario, P1 and P2 have all the same parametric values due to the fact that they both had the same initial belief, initial intent, and signal. Thus, walking through this scenario for P2 is the same as P1 described above.

Explanation of Data for Scenario 3

This scenario describes the Golden Ball example given earlier of an occasion when a player would lie about cooperating [26]. Similar to the description for scenario 2, the following is an explanation of all the different variables for player 1 (P1) and player 2 (P2) in scenario 3. Describing P1 first, they believe that P2 is going to cooperate (C). They signal D to P2, lying, because they intend to cooperate. Their COI is .6 since they plan to C, but P2 signals C, so there is a slight hope of getting the temptation payoff. CB is fixed at .8 meaning that they have a strong confirmation bias that P2 is going to be acting in a similar manner to them. P1's tells are low (.2) since P2 is telling the truth and not sending off many tells. Since they are lying, their sensitivity is low at .2. Gullibility and Fear are both held constant at .5 meaning P1 is in the middle of being gullible versus suspicious as well as having an average fear of getting the sucker's payoff. Equation 1 was utilized to determine the final S of P1 to be .6, the final CB to be .1 and the final COI to be .8. These three values were then averaged to get .5, which was compared to F (.5) to determine if P1 thinks P2 is telling the truth/lying (P1's credibility assessment/CA). Since avg1 = F, P1 thinks P2 is telling the truth, which is a correct (R) assessment (A). Thus, P1 believes P2 is going to play C, so P1 plays C. Due to having a COI of .6, the prefrontal lobe payoff (PF) of P1 is -30, since it is putting stress that is a little above average on P1's body. The stress placed coming from the limbic system (LS) is -50 since P1 is being deceitful. Since tells are low, the sensory information entering the parietal lobe (PA) is low yielding a -10 payoff. These three payoffs were added to get the total biological payoff of -90. The game payoff for P1 is 100 since the final play of moves was C, C meaning both players split the reward. The net payoff for this individual (the game plus the biological payoff) is 10, but this net value is only accumulated if the game and biological payoffs are in the same units.

Now describing P2, they believe that P1 is going to cooperate (C) along with signaling C to P2. They are telling the truth because they intend to cooperate. Their COI is .2 since P2 signals D, so there is almost no hope of getting the temptation payoff. CB is fixed at .8 meaning that they have a strong confirmation bias that P2 is going to be acting in a similar manner to them. P2's tells are high (.8) since P1 is lying and giving off many tells. Since P2 are telling the truth, their sensitivity is high at .8. Gullibility and Fear are both held constant at .5 meaning P1 is in the middle of being gullible versus suspicious as well as having an average fear of getting the sucker's payoff. Equation 1 was utilized to determine the final S of P1 to be 0, the final CB to be .9 and the final COI to be .4. These three values were then averaged to get .4, which was compared to F (.5) to determine if P2 thinks P1 is telling the truth/lying (P2's credibility assessment/CA). Since avg1 < F, P1 thinks P2 is telling lying, which is a correct (R) assessment (A). Thus, P2 believes P1 is going to play C, so P1 plays C. Due to having a COI of .2, the prefrontal lobe payoff (PF) of P1 is -10, since it is putting stress that is a little above average on P2's body. The stress placed coming from the limbic system (LS) is 0 since P2 is being truthful. Since tells are high, the sensory information entering the parietal lobe (PA) is high yielding a -40 payoff. These three payoffs were added to get the total biological payoff of -50. The game payoff for P2 is 100 since the final play of moves was C, C meaning both players split the reward. The net payoff for this individual (the game plus the biological payoff) is 50, but this net value is only accumulated if the game and biological payoffs are in the same units.

For ease of reference, the key below was developed containing the variables described above.

Symbol Quick Reference Key

Bi/f # = initial/final belief of player # (forecast of player ~j's move)

Sig # = signal of player #

I # = intent of player #

COI/COIf # = initial/final conflict of interest of player #

CB/CBf # = initial/final conflict of interest of player #

S/Sf # = initial/final sensitivity of player #

T # = tells of player #

G # = gullibility of player #

F # = fear of player #

Avg = average of S1f, CB1f and COI1f = $\frac{Sf+CBf+COIf}{3}$

CA # = credibility assessment of player #

A # = assessment validity of player #

MOVE # = move of player #

L = lie

TR = truth

R = right

W = wrong

C = cooperate

D = defect

π # = game payoff of player #

bio π = collateral biological payoff of player # = PF+LS+PA

net π # = sum of game and biological payoffs for player # = $\pi + bio\ \pi$

PF = payoff associated with the prefrontal cortex

LS = payoff associated with the limbic system

PA = payoff associated with the parietal lobe

Chapter 4

Model Analysis and Results

After taking desire out of the equation, all possible combinations of belief, intent, and signal were found. There are 4 various combinations per player, making a total of 16 variations taking both players into consideration. The initial and final beliefs of each player must be investigated due to the various parameters that factor into Equation 4 that impact each individual's final beliefs, thus correspondingly each player's move. The parameters feeding into the credibility assessment equation are conflict of interest (COI), confirmation bias (CB), tells (T), sensitivity (S), fear (F), and gullibility (G). These parameters will be analyzed on a 0-1 scale indicating the strength of the parameter, which can be endogenous or exogenous depending on the effect of player actions on parameter values. The endogenous variables considered are desire, initial intent, initial belief, G, T and F. The exogenous variables are the variables computed, which are final S, COI, and CB. As described in the previously, these psychological parameters compiled enumerate the individual's brain activity.

The model was constructed based on data gathered from a compilation of sources. The goal of the model is to display numerically the probabilities and likelihood that each player will have a correct detection of the other playing telling the truth/lying (analyzed by Equation 4). Once the truth/defection is detected, this information along with the biological significance of lying will be utilized to determine if lying is beneficial for each scenario.

Equation 4 describes the credibility of assessment of player not j (~j) by player j. The components of the equation are the intention (I) and honesty (h) of player j and ~j. The intention corresponding to player j is their intended move. For example, in the Prisoner's Dilemma, their intended move could be either to cooperate (C) or defect (D). Player j's honesty corresponds to whether or not player j is lying or telling the truth. This is whether the intention of player j matches the signal they send to player ~j. The parameters affecting these factors of the equation were determined by deciphering how the brain senses and processes opponents lying. Equation 4 will be analyzed for both players in the game to determine their credibility assessment (whether or not they think their opponent is lying or telling the truth).

A few of the variables were held constant since they would be altered based on the individual's personality. The model can be tested using different values held constant as well, yielding similar results. For this numerical example, CB was held constant at .8 being that the individuals have strong biases where they are looking to confirm that their beliefs of the opponent are true. G was held constant at .5 being the player has an average level of gullibility compared to suspicion. F was also held at .5 meaning that they obtain an average level of fear that the opponent is going to play defect and leave them to have the sucker's payoff. For each scenario described below, Table 3 was utilized for P1 and Table 4 for P2.

Factors of Player 1 and 2

Table 3: Data for Player 1

1	B1i	Sig1	Truth 1	I1	COI1	CB1	T1	S1	G1	F1	S1f	CB1f	COI1f	avg1	CA1	A1	B1f	MOVE 1
2	C	C	TRUE	C	0.6	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.8	0.8	TR	R	C	C
3	C	D	FALSE	C	0.6	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.8	0.5	TR	R	C	C
4	C	C	TRUE	C	0.2	0.8	0.8	0.8	0.5	0.5	0	0.9	0.4	0.4	L	R	C	C
5	C	D	FALSE	C	0.2	0.8	0.8	0.2	0.5	0.5	0	0.1	0.4	0.2	L	R	C	C
6	D	C	FALSE	D	0.9	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.95	0.6	TR	R	C	C
7	D	D	TRUE	D	0.9	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.95	0.8	TR	R	C	C
8	D	C	FALSE	D	0.2	0.8	0.8	0.2	0.5	0.5	0	0.1	0.4	0.2	L	R	C	C
9	D	D	TRUE	D	0.2	0.8	0.8	0.8	0.5	0.5	0	0.9	0.4	0.4	L	R	C	C
10	C	C	TRUE	C	0.6	0.8	0.8	0.8	0.5	0.5	0	0.9	0.8	0.6	TR	W	C	C
11	C	D	FALSE	C	0.6	0.8	0.8	0.2	0.5	0.5	0	0.1	0.8	0.3	L	R	D	D
12	C	C	TRUE	C	0.2	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.4	0.6	TR	R	D	D
13	C	D	FALSE	C	0.2	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.4	0.4	L	R	C	C
14	D	C	FALSE	D	0.9	0.8	0.8	0.2	0.5	0.5	0	0.1	0.95	0.4	L	R	D	D
15	D	D	TRUE	D	0.9	0.8	0.8	0.8	0.5	0.5	0	0.9	0.95	0.6	TR	W	C	C
16	D	C	FALSE	D	0.2	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.4	0.4	L	W	C	C
17	D	D	TRUE	D	0.2	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.4	0.6	TR	R	D	D

Table 3 is half of the model, which contains player 1's data. The numerical values for the 6 parameters were generated based on the specifications described in the above section. When calculating the final COI, CB, and S values, G represented the initial starting value (which can be seen in equations 1, 2, and 3). Fear was compared to the average (avg) values, and if the avg value was greater than or equal to the F value then the player believes the opponent is telling the truth, and if the value is below this threshold value then the player believes the opponent is lying, which is represented in the CA column. Depending on player j's detection is whether or not they correctly assessed player ~j's move (represented as A). Based on if player j thought player ~j was telling the truth or not dictated the final belief of player j and ultimately player j's move.

Table 4: Data for Player 2

1	B2i	Sig2	Truth 2	I2	COI2	CB2	T2	S2	G2	F2	S2f	CB2f	COI2f	avg2	CA2	A2	B2f	MOVE 2
2	C	C	TRUE	C	0.6	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.8	0.8	TR	R	C	C
3	C	C	TRUE	C	0.2	0.8	0.8	0.8	0.5	0.5	0	0.9	0.4	0.4	L	R	C	C
4	C	D	FALSE	C	0.6	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.8	0.5	TR	R	C	C
5	C	D	FALSE	C	0.2	0.8	0.8	0.2	0.5	0.5	0	0.1	0.4	0.2	L	R	C	C
6	C	C	TRUE	C	0.6	0.8	0.8	0.8	0.5	0.5	0	0.9	0.8	0.6	TR	W	C	C
7	C	C	TRUE	C	0.2	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.4	0.6	TR	R	D	D
8	C	D	FALSE	C	0.6	0.8	0.8	0.2	0.5	0.5	0	0.1	0.8	0.3	L	R	D	D
9	C	D	FALSE	C	0.2	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.4	0.4	L	W	C	C
10	D	C	FALSE	D	0.9	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.95	0.6	TR	R	C	C
11	D	C	FALSE	D	0.2	0.8	0.8	0.2	0.5	0.5	0	0.1	0.4	0.2	L	R	C	C
12	D	D	TRUE	D	0.9	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.95	0.8	TR	R	C	C
13	D	D	TRUE	D	0.2	0.8	0.8	0.8	0.5	0.5	0	0.9	0.4	0.4	L	R	C	C
14	D	C	FALSE	D	0.9	0.8	0.8	0.2	0.5	0.5	0	0.1	0.95	0.4	L	R	D	D
15	D	C	FALSE	D	0.2	0.8	0.2	0.2	0.5	0.5	0.6	0.1	0.4	0.4	L	W	C	C
16	D	D	TRUE	D	0.9	0.8	0.8	0.8	0.5	0.5	0	0.9	0.95	0.6	TR	W	C	C
17	D	D	TRUE	D	0.2	0.8	0.2	0.8	0.5	0.5	0.6	0.9	0.4	0.6	TR	R	D	D

Table 6 mimics the structure of Table 5; however, the computations are all for player 2. The columns here correspond to the columns of Table 5, for each scenario of play.

The three equations incorporated in Table 3 and 4 were constructed to feed into the calculation of Equation 4, which calculates if the player correctly assesses the other player's move. Correct assessment is based on if the player has a correct final belief, which is determined based on taking the average of the final sensitivity, confirmation bias, and conflict of interest of the player. The average of equation 1, 2, and 3 is calculated to result in the assessment of player j. The average is $\frac{Sf+CBf+COIf}{3}$. Based on the assessment, player j chooses their move. Each scenario (2-17) will now be described.

2. Player 1: P1 obtains the initial belief that P2 is going to cooperate. Signals to P2 that P1 is going to cooperate and is telling the truth because P1's intent is also to cooperate. P1 has an above average COI of .6, since P2 signals C, so P1 is tempted to play D; however, not fully tempted since P1 intends to cooperate as well. T is held at .2 since the other player is telling the truth. S is .8 since P1 is telling the truth and their parietal lobe's activity is not dedicated to covering up that they are lying. The final calculation of S is .6, CB is .9, and COI is .8 based on the calculations performed. The average of these three came to be .8, which is higher than F of .5, so P1 detects that P2 is telling the truth. Since P2 is telling the truth this is an accurate assessment. Due to this, the final B is that P2 is going to cooperate, so P1 plays C as well.

Player 2: For this scenario, P2 has all the same data as P1.

3. Player 1: P1 obtains the initial belief that P2 is going to cooperate. Signals to P2 that P1 is going to defect and is lying because P1's intent is to cooperate. P1 has an above average COI of .6, since P2 signals C, so P1 is tempted to play D; however, not fully tempted since P1 intends to cooperate as well. T is.2 since the other player is telling the truth. S is .2 since P1 is lying and their parietal lobe's activity is mainly focused on P2 not detecting their lie. The final calculation of S is .6, CB is .1, and COI is .8 based on the calculations performed. The average of these three came to be .5, which is equal to F of .5, so P1 detects that P2 is telling the truth. Since P2 is telling the truth this is an accurate assessment. Due to this, the final B is that P2 is going to cooperate, so P1 plays C as well.

Player 2: P2 believes P1 is going to cooperate. P2 intends to cooperate and signals cooperate, so they are telling the truth. P2 has a low COI of .2 since P1 signals D, so they do not have hopes of getting the temptation payoff. T is .8 since P1 is lying and is giving off lots of tells that they are lying. S is .8 since P2 is telling the truth and has the ability to be observant. The final S is 0, CB is .9, and COI is .4.

The average of these values is .4, so P2 determines P1 is lying, which is the right assessment. Thus, P2 has a final belief of C, so makes the move of C.

4. Due to symmetry with row 3, refer to row 3's description; however, P1 and P2's descriptions are reversed.

5. Player 1: P1 suspects P2 is going to cooperate. Although P1 intends to play C, they signal D, meaning that they are lying to P2. Their COI is .2 since P2 signals defect, so they have little hope of getting the temptation payoff. T is .8 and S is .2 since both players are lying. S1f is 0, CBf is .1, and COIf is .4 making the average of the three .2. Thus, P1 detects P2 is lying, which gives a right assessment. B1f is C, so P1 also cooperates for their move.

Player 2: For this scenario, there is symmetry between P1 and P2, so the results of P2 are the same as P1.

6. Player 1: Player 1 is intending to play D, but is lying, so signals C. Since his intent is D, he projects that view onto his belief, believing P2 is intends to play D too. Since P2 signals C and P1 intends to play D, the temptation payoff is extremely tempting making COI .9. T and S are both .2 since P2 is telling the truth and P1 is lying. S1f is .6, CB1f is .1 and COIf is .95 averaging to .6. This average is greater than F, so P1 determines that P2 is telling the truth, which is an accurate assessment. Thus, P1 has the final belief that P2 is going to cooperate, so P1 also plays C.

Player 2: P2 is telling the truth that they are going to cooperate. Since they intend to cooperate, they believe P1 is going to cooperate too. Their COI is .6 since P1 signals C, so there is some temptation to play D and get the temptation payoff. T and S are both .8 since P1 is lying and P2 is telling the truth. S1f is 0, CBf is .9 and COIf is .8, averaging to .6. Hence, P2 determines that P1 is telling the truth; however, this is an incorrect assessment. In the end, P2 believes P1 is going to cooperate, so P2 cooperates.

7. Player 1: P1 intends to defect, so they believe that P2 does too. P1 is telling the truth and signals D. Since P2 signals C and P2 intends to play D, the temptation payoff is extremely tempting making COI .9. T is .2 and S is .8 since P1 and P2 are both telling the truth. S1f is .6, CB1f is .9 and COIf is .95, which average to .8 meaning that P1 detects that P2 is being truthful, which is a correct assessment. P1 finalizes their belief as C, so moves C.

Player 2: P2 is telling the truth about cooperating. They believe P1 is also going to cooperate. Their COI is .2 since P1 signals D, so their hopes of the temptation payoff are low. T is .2 and S is .8 since both players are telling the truth. Post calculations, Sf is .6, CBf is .9 and COIf is .4. These values average to .6 leading P2 to believe P1 is telling the truth, which is right. Thus, P2 believes P1 is going to defect, so P2 defects as well.

8. Player 1: Intending to defect, P1 signals C, lying to P2. COI is .2 since P2 signals D as well. T is .8 and S is .2 since both players are lying. After inserting the respective values into Equation 2, 3 and 4, Sf is 0, CBf is .1 ad COIf is .4. These average to be .2, leading P1 to believe P2 is lying about defecting. Due to this, P1 believes P2 is going to cooperate, so P1 also cooperates. P1’s assessment of P2 was correct.

Player 2: P2 is also lying, but intends to play C and signals D, unlike P1. P2 believes P1 is going to play C. Since P1 signals D, P2 has a COI of .2. T and S are .8 and .2, respectively. Sf is 0, CBf is .1 and COIf is .8, averaging to .3 leading P2 to think P1 is lying, which is false. P2 believes P1 is going to play C, so P2 plays C as well.

9. This scenario is symmetric to scenario 7. All of the data is the same, but just switch P1 of scenario 7 to be P2.

10. Scenario 6 is the reverse of this scenario, so refer to scenario 6; however, switch P1 and P2.

11. Scenario 8 is the reverse of this scenario, so refer to scenario 6; however, switch P1 and P2.

12. Scenario 7 is the reverse of this scenario, so refer to scenario 6; however, switch P1 and P2.

13. Scenario 9 is the reverse of this scenario, so refer to scenario 6; however, switch P1 and P2.

14. Player 1: Player 1 is intending to play D, but is lying, so signals C. Since his intent is D, he projects that view onto his belief, believing P2 is intends to play D too. Since P2 signals C and P1 intends to play D, the temptation payoff is extremely tempting making COI .9. T is .8 and S is.2 since P1 is lying and P2 is telling the truth. S1f is 0, CB1f is .1 and COIf is .95 averaging to .4. The average is less than F, so P1 determines that P2 is telling lying, which is an accurate assessment. Thus, P1 has the final belief that P2 is going to defect, so P1 also plays D.

Player 2: P2 has all of the same values as P1 for this scenario.

15. Player 1: P1 intends to defect, so they believe that P2 does too. P1 is telling the truth and signals D. Since P2 signals C and P2 intends to play D, the temptation payoff is extremely tempting making COI .9. T and S are both .8 since P2 is lying and P1 is telling the truth. S1f is 0, CB1f is .9 and COIf is .95, which average to .6 meaning that P1 decides that P2 is being truthful, which is a incorrect assessment. P1 finalizes their belief as C, so moves C.

Player 2: Player 2 is intending to play D, but is lying, so signals C. Since his intent is D, he projects that view onto his belief, believing P2 is intends to play D too. Since P2 signals D and P1 intends to play D, the temptation payoff seems out of reach, so COI .2. T and S are .2 since P1 is telling the truth and P2 is lying. S1f is .6, CB1f is .1 and COIf is .4 averaging to .4. This value is less than F, so P2 determines that P1 is telling lying, which is an inaccurate assessment. Thus, P2 has the final belief that P1 is going to cooperate, so P2 also plays C.

16. This scenario is symmetric to scenario 15. All of the data is the same, but just switch P1 of scenario 15 to be P2.

17. Player 1: P1 is telling the truth about defecting, so they also believe that P2 is going to defect. COI is .2 since P2 also signals d. T is .2 since P2 is telling the truth and S is .8 since P1 is telling the truth. S1f is .6, CBf is .9, COIf is .4, so the average of the three is .6. Being greater than the fear value, P1 determines P2 is telling the truth, which is correct. In the end, P1 believes P2 is going to play D, so they play D as well.

Player 2: For this scenario, there is symmetry between P1 and P2, so the results of P2 are the same as P1.

Table 5: Payoff Matrix Corresponding to the Model

		Player 2	
Player 1		Cooperate	Defect
	Cooperate	100, 100	-50, 200
	Defect	200, -50	0, 0

Table 5 is the payoff matrix associated with this model and the set payoff values.

Table 6: Players' Truth, Final Moves and Designated Game Payoffs

1	Truth 1	Truth 2	MOVE 1	MOVE 2	π1	π2
2	TRUE	TRUE	C	C	100	100
3	FALSE	TRUE	C	C	100	100
4	TRUE	FALSE	C	C	100	100
5	FALSE	FALSE	C	C	100	100
6	FALSE	TRUE	C	C	100	100
7	TRUE	TRUE	C	D	-50	200
8	FALSE	FALSE	C	D	-50	200
9	TRUE	FALSE	C	C	100	100
10	TRUE	FALSE	C	C	100	100
11	FALSE	FALSE	D	C	200	-50
12	TRUE	TRUE	D	C	200	-50
13	FALSE	TRUE	C	C	100	100
14	FALSE	FALSE	D	D	0	0
15	TRUE	FALSE	C	C	100	100
16	FALSE	TRUE	C	C	100	100
17	TRUE	TRUE	D	D	0	0

This table incorporates the players' truthfulness, their moves (which were determined based on the calculations in Tables 3 ad 4) and the payoffs corresponding to the final combination of moves. The payoff values were randomly generated based off the typical Prisoner's Dilemma payoff matrix.

Table 6 shows the final moves of each player and the various payoff values are dictated by the combination of moves. When both players cooperate, they split the reward of the arbitrarily chosen value of 200 and both have a net positive game payoff of 100. The first five scenarios, as well as scenario 9, 10, 13, 15, and 16 yield these payoffs. When one player cooperates, and the other player defects the player that cooperates receives the sucker payoff of -50. The player that defects wins the temptation payoff of 200. The final combination of moves is both players playing D. When they both play D, they each get a payoff of 0. These values are also displayed in Table 5, where player 1's payoff is listed first in the corresponding matrix coordinate. Player one is the row player and player two is the column player. For example, the top right quadrant of the payoff matrix can be read as P1 cooperates getting the payoff of -50 and P2 defects receiving the payoff of 200.

Table 7: Payoff Values

1	PF 1	LS 1	PA 1	bio π1	π1	net π1	PF 2	LS 2	PA 2	bio π2	π2	net π2
2	-30	0	-10	-40	100	60	-30	0	-10	-40	100	60
3	-30	-50	-10	-90	100	10	-10	0	-50	-60	100	40
4	-10	0	-50	-60	100	40	-30	-50	-10	-90	100	10
5	-10	-50	-50	-110	100	-10	-10	-50	-50	-110	100	-10
6	-50	-50	-10	-110	100	-10	-30	0	-50	-80	100	20
7	-50	0	-10	-60	-50	-110	-10	0	-10	-20	200	180
8	-10	-50	-50	-110	-50	-160	-30	-50	-50	-130	200	70
9	-10	0	-50	-60	100	40	-10	-50	-10	-70	100	30
10	-30	0	-50	-80	100	20	-50	-50	-10	-110	100	-10
11	-30	-50	-50	-130	200	70	-10	-50	-50	-110	-50	-160
12	-10	0	-10	-20	200	180	-50	0	-10	-60	-50	-110
13	-10	-50	-10	-70	100	30	-10	0	-50	-60	100	40
14	-50	-50	-50	-150	0	-150	-50	-50	-50	-150	0	-150
15	-50	0	-50	-100	100	0	-10	-50	-10	-70	100	30
16	-10	-50	-10	-70	100	30	-50	0	-50	-100	100	0
17	-10	0	-10	-20	0	-20	-10	0	-10	-20	0	-20

Due to symmetry, for analysis consider only the cases for player 1. The scenarios where individuals are lying are marked in red.

When lying, there is stress put on the body from the limbic system since the player has pressure from lying causing them to be stressed that the other player is going to detect their lie (-50 payoff for LS). When being truthful, there is no stress placed on the individual (0 payoff for LS). Regardless of truthfulness/deceitfulness, the prefrontal cortex and parietal lobe can contribute to the individual's stress. The prefrontal cortex's activity is based on F, G, CB, and COI; however, since all of these values are held constant except COI, we will focus on COI. The level of stress contribution to the body is determined based on the weight of COI. A COI of .2 corresponds to a -10 payoff, .6 corresponds to a -30 payoff, and the most stress (-45 payoff) is associated with a COI of .9. The activity of the parietal lobe is based on tells. The parietal lobe is most active when tells are high (.8) since it is processing a lot of sensory information placing stress on the individual for an accurate perception of these tells (-40 payoff for PA).

Figure 2: Game Payoff vs. Biological Payoff for Player 1 with Player 2's Initial Intent=C and Being Truthful

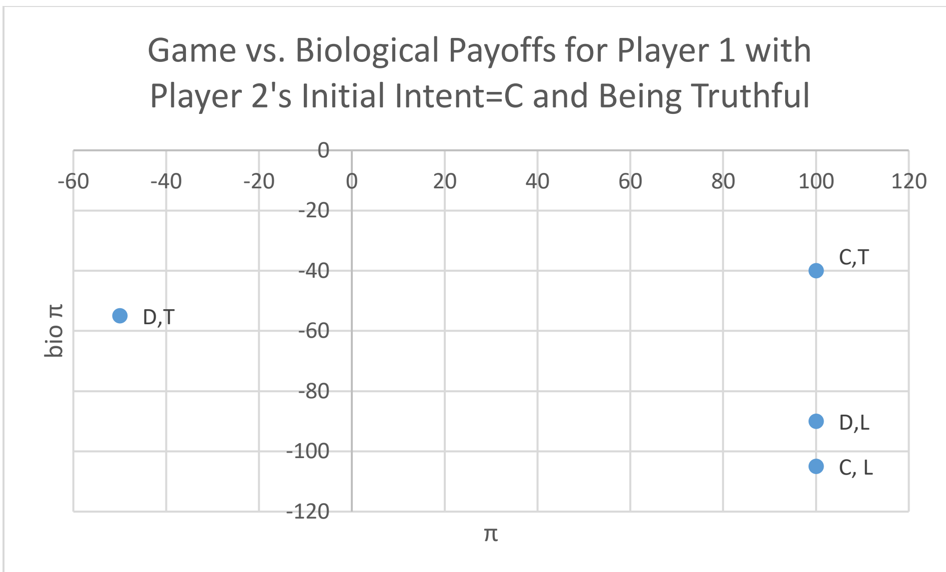

Figure 2 displays the game payoffs (π/ the x-axis variable) in comparison with the biological payoffs (bio π/ the y-axis variable) for player 1 when player 2 has the initial intent of cooperating (C) and is telling the truth. The points are labeled as player 1's initial intent (C/D) and whether they are telling the truth or lying (T/L). The Pareto boundary (the simultaneous optimization to try to maximize both payoffs) is a maximum in Figure 2, obtained at the point representing the scenario where player 1 has the initial intent of cooperating and is telling the truth.

Figure 3: Game Payoff vs. Biological Payoff for Player 1 with Player 2's Initial Intent=C and Lying

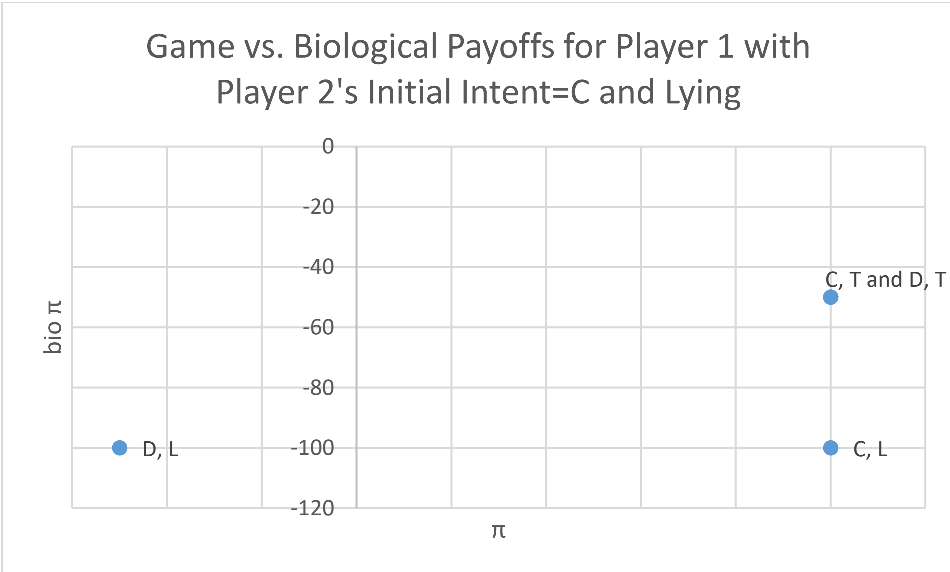

Figure 3 displays the game payoffs (π/ the x-axis variable) in comparison with the biological payoffs (bio π/ the y-axis variable) for player 1 when player 2 has the initial intent of cooperating (C) and is lying. The points are labeled as player 1's initial intent (C/D) and whether they are telling the truth or lying (T/L). The Pareto boundary (the simultaneous optimization to try to maximize both payoffs)) in Figure 3 is the point representing the scenarios where player 1 has the initial intent of cooperating or defecting and is telling the truth.

Figure 4: Game Payoff vs. Biological Payoff for Player 1 with Player 2's Initial Intent=D and Being Truthful

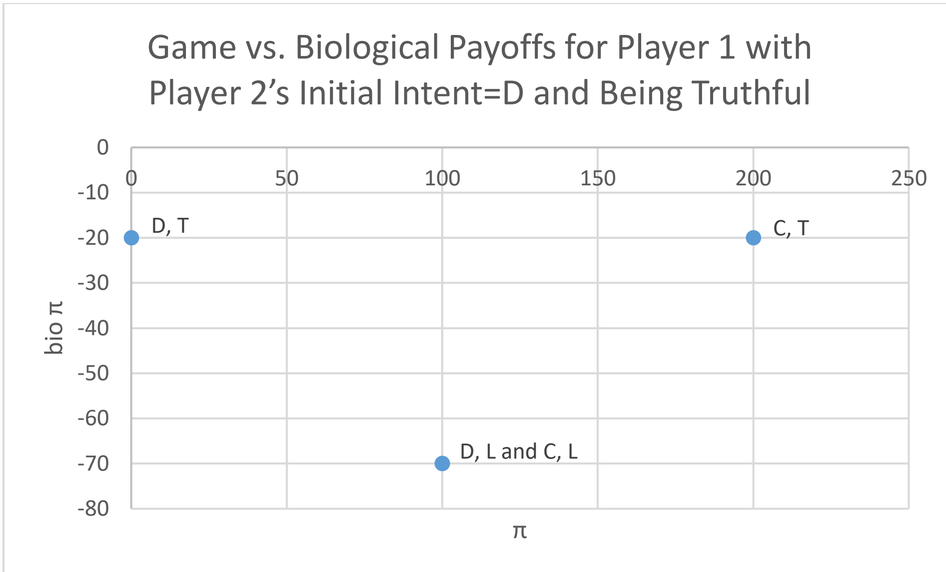

Figure 4 displays the game payoffs (π/ the x-axis variable) in comparison with the biological payoffs (bio π/ the y-axis variable) for player 1 when player 2 has the initial intent of defecting (D) and is telling the truth. The points are labeled as player 1's initial intent (C/D) and whether they are telling the truth or lying (T/L). The Pareto boundary (the simultaneous optimization to try to maximize both payoffs)) in Figure 4 is the point representing the scenario where player 1 has the initial intent of cooperating and is telling the truth (the same as seen in Figure 2).

Figure 5: Game Payoff vs. Biological Payoff for Player 1 with Player 2's Initial Intent=D and Lying

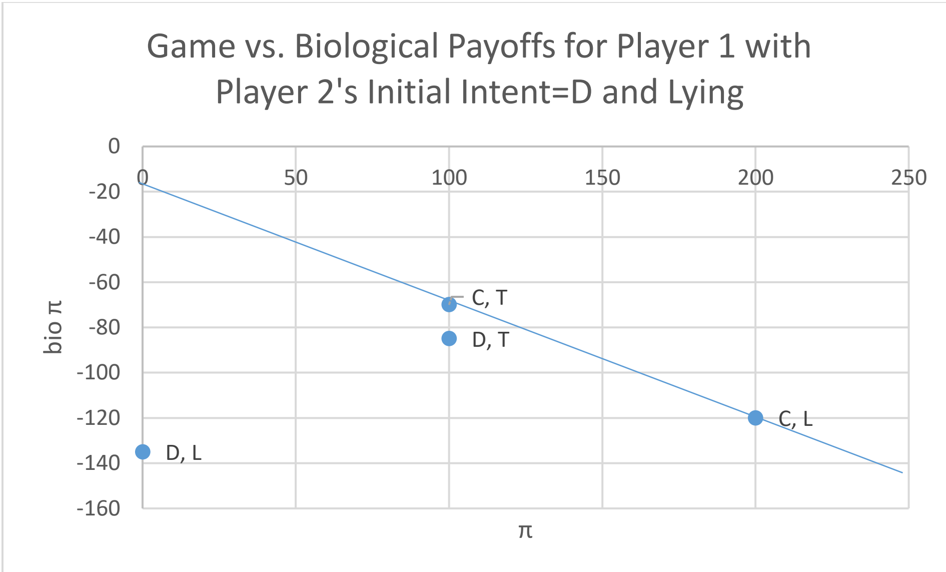

Figure 5 displays the game payoffs (π/ the x-axis variable) in comparison with the biological payoffs (bio π/ the y-axis variable) for player 1 when player 2 has the initial intent of defecting (D) and is lying. The points are labeled as player 1's initial intent (C/D) and whether they are telling the truth or lying (T/L). The Pareto boundary (the simultaneous optimization to try to maximize both payoffs) in Figure 5 is the line connecting the points that represent the scenarios where player 1 has the initial intent of cooperating and is either telling the truth or lying (the blue line). This line represents the trade-off of maximizing one payoff, but lowering the other, so the trade-off of only being able to maximize one payoff in this scenario is shown.

Figures 2-5 display the trade-off spaces for the game and biological payoffs for each scenario displayed in Tables 3 and 4. To see clearly the best strategy for player 1 depending on player 2's strategy, the scenarios were plotted based on player 2's initial conditions of intent and truthfulness. For maximizing the trade-off space between the game and biological payoffs, both are plotted for each scenario to determine the Pareto boundary (the boundary where past this there is no way to optimize both payoffs involved). This strategy was utilized because the game and biological payoffs may not be in the same units leading to the trade-off space mentioned.

Chapter 5

Discussion

The model ended up being a truth detector as well as a median to determine if lying is beneficial in certain scenarios. Answering the overarching question of this research, the final payoffs of each player in the various scenarios along with the neurological activity for each scenario must be considered to compile the full biological effect on player j. The payoffs used here are for demonstration purposes only to establish a structure that can be validated with more experimentation.

Starting with the game payoff values each player gets, Table 6 displays that the majority of the time (10/16 times) an individual lies, they receive a payoff of 100. Two out of the sixteen occurrences where a player lies, they receive the maximum payoff of 200. Two instances out of the sixteen times individuals lie in the model they end up with a payoff of zero. 2/16 of the lying scenarios, player j receives a negative payoff of -50. Therefore, 87.5% of the time an individual lied they received a net positive payoff. The same ratios are seen for the truthful players as well, so 87.5% of the time player j tells the truth they receive a positive payoff.

When lying, player j is fundamentally disconnecting them self from reality, which can be a benefit in the form of false optimism, but it can also be a liability reducing their sensitivity to player ~j. A confidence factor could be a benefit or not depending on if it gives players the faith to believe in themselves. This factors into the final payoff of player j through sensitivity.

The ability to alter intent into the final move is based on whether player j thinks player ~j is lying. For example, if player j lies then they are disconnected from reality and do not detect that player ~j is also lying about cooperating, so player j does not change their final move and cooperates, but player ~j defects leaving player j with the sucker payoff. In this scenario, it is not worth it to lie because player j gets harmed in the end.

Lying when signaling D and intending to play C could be argued nonsensical; however, it can end up being beneficial as seen in scenario 3 for player 1 (in Table 3). It was also displayed in the Golden Ball gameshow where player j signaled D but intended to play C and the other player either detected this lie, or just accepted that in order to have a chance of receiving any of the prize they would have to play C [26]. Player j lying clouded their ability of a correct assessment which could have ended up being a tangible negative. Player ~j would have had a better payoff if they successfully ripped the other player off and lied about it; however, in a repeated game this hurts player ~j because player j will no longer trust ~j in the beginning and is more likely to defect against ~j.

Understanding the psychological functioning of lying is crucial for determining the biological stress placed on player j's body. The adjoining of activity, whether it be a positive or negative contribution to the immune function, will be defined as the biological payoff. Summing the biological payoff with the game payoff (when they are in the same units, such as evolutionary fitness) for player j will instill the net payoff of player j. The net payoff will deduce whether lying was truly beneficial to player j when the biological and game payoffs are in the same units.

With this model exemplifying the first time an individual lies, the stress player j goes through when they lie is the highest they will ever experience. As mentioned in the introduction, stress takes a toll on the body in a variety of ways, but the primary way we will focus on is depletion of the immune function. This depletion impacts every physiological system by putting the body at risk of being attacked by a foreign antigen along with making these systems less functional causing premature ageing. Shame and guilt from lying depress the immune function and an individual's overall health. Repressing the truth from the conscious mind also lowers the immune function and brain activity since repression takes a lot of energy, so even if one is not shameful about lying, they still are depressing their immune function by lying. The mere suppression of information also decreases immune system ability. Pertaining to people having a strong confirmation bias, they suppress information that does not support their bias so that their bias is strengthened even without adding supporting evidence. Shown in Tables 3 and 4, a strong confirmation bias of .8 was held constant contributing a negative payoff value due to its harmful health effects.

Along with CB contributing to activity in the prefrontal cortex, F, G, and COI also are active in this lobe. Since F and G were held constant at .5, they yield an average activity in this lobe. With CB and COI being the main factors dictating the overall level of activity in this lobe, COI only adds onto the increased activity from CB. COI varying from .2, .6, and .9, individuals who lied and obtained a COI of .9 have the highest level of activity possible in this lobe. On the other hand, activity is not necessarily a hindrance. If COI is high and the player does not end up getting a positive payoff, then it is a hindrance because they got their hopes up of getting the temptation payoff. Observing the results from Table 3, P1 from scenario 6 and 14 had the highest COI possible and were lying, but only P1 from 6 received a payoff of 100, while P1 from 14 received a payoff of 0. This is similarly seen in Table 4 with P2 for scenarios 10 and 14. The only players who ended up winning the temptation payoff of 200 were P1 for scenarios 11 and 12, and P2 for scenarios 7 and 8. For scenario 11/8, P1/P2 was lying about cooperating and had a moderate COI of .6. For scenario 12/7, P1/P2 was telling the truth about cooperating and had a low COI of .2. Hence, a high COI never yields the temptation payoff (with the parametric values set in Tables 3 and 4), so the higher the COI the more detracting to the player's biological impact.

The limbic system is the main portion of the brain that controls stress. A player's sensitivity, fear and gullibility all factor into the activity of the limbic system. With F and G being held constant, we will focus on S. For S to be low is more detrimental to the limbic system since having a depressed ability to sense player ~j's lie along with player j trying to cover up their own lie is stressful. All instances of lying, player j has a S of .2. Thus, every time a lie occurs the payoff for the limbic system is negative.

Corresponding to sensitivity is tells. T deals with activity in the parietal lobe since player j is observing the actions and tells of player ~j. Whether they are detected or not are based on sensitivity. Therefore, S multiplied by T determines the level of brain function in this lobe. Low sensitivity and high tells generates a negative implication on the payoff for the parietal lobe. Even though vice versa high sensitivity and low tells promotes the same product, this does not cause stress on the body. From Table 3, when lying, P1 has a S of .2 and T of .8 for scenarios 5, 8, 11, and 14. From Table 4, lying P2 has a S of .2 and T of .8 in scenarios 5, 8, 11 and 14 as well.

The scenarios where player j lies in Table 7 are displayed in red. The activity in each brain location ranges from -50 to 0 with -50 being activity causing the highest possible stress on player j's body and 0 meaning the activity in this region has no impact on player j's physiological status.

As shown in Table 7, 4/8 times an individual lies they receive a net negative payoff of -10, -100 or -160, 4/8 times they receive a net positive payoff of 10, 20 or 30. This means approximately 50% of the time an individual lies it is beneficial to them. The 50% of scenarios with a net negative payoff, the net negative payoff is more significant of an impact than the net positive payoff. To calculate expected value (the weighted average) for a data set, one must take each value and weight it by the probability of occurrence = 1/total case number where p1 lies and add them all up, which gives a weighted average that is the expected benefit of lying. When the same units are considered (for example, the game and biological payoffs both being in the units of evolutionary fitness), the expected value (E[x]) for scenarios when the individual lies is the following:

$$E(x) = \frac{1}{8} * 10 + \frac{1}{8} * (-5) + \frac{1}{8} * (-150) + \frac{1}{8} * 80 + \frac{1}{4} * 30 + \frac{1}{8} * (-135) = -17.5.$$

On the contrary, only 1/8 of the scenarios where the individual tells the truth do, they receive a net negative payoff. This scenario for P1 is scenario 8 where the individual receives the sucker's payoff. The expected value for scenarios when the individual tells the truth is the following:

$$E(x) = \frac{1}{8} * 60 + \frac{1}{4} * 50 + \frac{1}{8} * (-105) + \frac{1}{8} * 30 + \frac{1}{8} * 180 + \frac{1}{8} * 15 + \frac{1}{8} * (-20) = 32.5.$$

Thus, when considering game and biological payoffs in the same units, on average, it is not beneficial to lie.

When the game and biological payoffs are in different units (such as the game payoff being a physical reward versus the biological payoff being immune deficiency), the separate expected values for the game and biological payoffs of individuals lying versus telling the truth were also considered. For individuals lying, the expected value of the game payoff is as follows.

$E(x) = \frac{5}{8} * 100 + \frac{1}{8} * (-50) + \frac{1}{8} * 200 = 81.25$. The expected value for the biological payoff for individuals lying is:

$$E(x) = \frac{1}{8} * (-90) + \frac{1}{4} * (-100) + \frac{1}{8} * (-105) + \frac{1}{8} * (-120) + \frac{1}{4} * (-70) + \frac{1}{8} * (-135) = -98.75.$$

Consequently, in this case, the mean game payoff is a net positive value of 81.25, but the biological positive is a net negative value of -98.75. Furthermore, the biological payoff on average is has a greater negative effect on the player than the average positive game payoff the individual can receive.

Splitting the payoff values into the varying unit categories, for an individual telling the truth the expected value for the game payoff is the same as for an individual who lies.

$E(x) = \frac{5}{8} * 100 + \frac{1}{8} * (-50) + \frac{1}{8} * 200 = 81.25$. However, the expected value of the biological payoff for the truthful individual is much lower than for the deceitful average.

$E(x) = \frac{1}{8} * (-40) + \frac{1}{4} * (-50) + \frac{1}{8} * (-55) + \frac{1}{8} * (-70) + \frac{1}{4} * (-20) + \frac{1}{8} * (-85) = -48.75.$

Hence, the mean negative biological payoff does not outweigh the mean positive game payoff.

Besides expected value, one must observe the trade-off space shown in Figures 2-5 to determine if the physical payoff is worth the collateral harm. The area under the Pareto boundary is where the individual can shift their biological and game payoffs to try to maximize both. Once the Pareto boundary is met, the player cannot increase one payoff without the other one getting worse. The game payoff (the horizontal/x-axis) increases moving to the right and the biological payoff (the vertical/y-axis) decreases moving down the axis/further South.

As seen in Figure 2, when player 2 has the initial intent to cooperate and is telling the truth, the maximum payoff combination possible for player 1 (biological payoff of -40 and game payoff of 100) is achieved when player 1 intends on cooperating and is telling the truth. This is also the case for when player 2 is intending to defect and is telling the truth with player 1 receiving a biological payoff of -20 and a game payoff of 200 (Figure 4). Figure 3 confirms that player 1 receives their maximum payoff combination whey they are truthful, and for player 2 intending to cooperate yet lying, truthfulness is player 1's best strategy regardless of their intent of play. The Pareto boundary for Figure 5 is not a point like the other figures, but a line where the trade-off between maximizing biological versus game payoffs can be seen. If a player wants the maximum game payoff possible for the scenario where player 2 intends to defect and is lying, they either have to choose a better game or biological payoff, they cannot have the maximum possible value for both. Hence, they can choose a larger game payoff of 200 that comes with the larger biological harm of -120, or they can pick the larger biological payoff of -70 that comes with the lesser game payoff of 100. Here it is up to the player's discretion to determine if lying or telling the truth is more beneficial to them; however, in the prior three figures, the optimal strategy where the player benefits the most is by telling the truth.

Overall, for both types of scenarios (where the units of the game and biological payoffs are the same or different) based on the expected values calculated and observing the Pareto boundaries for all scenarios, it is not beneficial to lie. Different values of T, R, S and P were tested under the condition that T > R > P > S, and the same result (that lying is never truly beneficial) was found; however, this model must still be validated.

Chapter 6

Future research

Repeated Game

This model could be altered to be a repeated game where players can encounter each other to play again after the original scenario. In the repeated game, updated parameters would need to come into play since there is now an added dimension of the players being familiar with the other's history of play. Past tendencies of player ~j will now factor into player j's initial belief of what player ~j is going to do. Take for example a two-period game where in period 1 player j successfully ripped player ~j off by defecting when player ~j cooperated. In period two, player ~j is less likely to trust that player j is telling the truth, defecting in the end so that player ~j does not receive the sucker payoff again.

With being a repeated game, the activity levels of the various brain parts are changed as well. There is an inverted relationship between times an individual lies and that individuals stress level when lying, which has the overall contribution to the individual's biological payoff. As a player lies more frequently then their amygdala becomes less active, decreasing the amount of stress placed on the body not affecting the immunological function of the player. Hence, in a repeated game it is more likely that lying evolves to be beneficial since the biological payoff does not negate from the net payoff.

Adding Desire and Personality

Desire was held constant at reward for this study. In the future, desire could be added back so that the player has the option of hoping for either the reward, temptation, or punishment payoff. Adding this extra dimension would allow for each player to have an initial and final desire. Since desire is a starting parameter that motivates player j's intent and player ~j's signal shapes player j's final intent, player j's desire could also be updated. Particularly due to the fact that desire and expectation do not have to match. This could also factor into the future repeated game model where the desire of a player can vary throughout the time sequence of periods.

Along with varying desire, individuals' personalities can be considered to contribute to the various parameters. Personality would have a large impact on most of the parameters. It would alter COI depending on if the player is optimistic or cares about getting the temptation payoff. It would affect CB depending on how confident and trusting the player is. T could be changed if a player's personality affects how they convey tells. If player j is a typically nervous individual, then they could be giving off tells that they are lying even when they are not. With this same example, if player j is lying then their tells would be even higher in comparison to another individual since their nervousness is extreme. Based on personality, S can vary if the individual tends to be observant or inattentive. If inattentive and lackadaisical then the players sensitivity would be lowered, whereas if a player is observant and mindful then their sensitivity would rise. G would obviously change if the person is more gullible or skeptical. Along with F which would adjust based on if the player is a fearless or fearful person.

With repeated games this is an interesting addition because an individual's personality can change over time. Also, player j being accustomed to player ~j's personality due to repeated encounters can have a large weight on how the game plays out. If player j does not detect player ~j's personality digression, then player j potentially would be less likely to have a correct assessment.

Analyzing Various Scenarios

At the beginning of this study, a plethora of scenarios were thought of that could impact the various parameters for each player. Some examples of social situations in which people lie are listed below. These social situations could be explored further by altering the model to fit a particular situation.

1. Interpersonal relationships
    a. Self-Deception
        i. Give confidence
        ii. Motivate
        iii. To believe something that isn't true
        iv. Social advantage
        v. Comfort yourself
        vi. Avoid change
        vii. Avoid hard questions
            1. According to research done by Leonard Pitts, Jr., people that pay more for forgeries are less likely to admit it is a forgery
            2. Staying consistent about a lie
            3. Live with lie then exposed to reality
    b. Family
        i. Parents lying to child to "protect them"
            1. Disease/death

        2. Financial stability
        3. Family drama
        4. Lie to spouse about cheating
    ii. Children lying to parents to not "worry" them
        1. Grades
        2. Health
        3. Personal problems (friend drama/living situation/relationships)
        4. Sibling drama
        5. Get what they want
    iii. Relatives lying
        1. Grandparents lying to grandchildren to allude that they were always well behaved to set a good example
        2. Lying to siblings to get what you want
        3. Lying to cousins/aunts/uncles to seem like your life is together

c. Friends
    i. Lying to go where you want to go
    ii. Lying to get something you want (deter them from wanting it)
    iii. Lying because you did something that you think will hurt them
    iv. Sugar coating things
    v. Trying to maintain a certain appearance
    vi. Make yourself seem more appealing

d. Coworkers/peers
    i. lying to get ahead in the office/classroom
    ii. lying to professor to get an extension/better grade
    iii. lying to get a raise
    iv. lying to get information/notes
    v. lying to not go to class

e. Police
    i. get out of a charge/ticket
    ii. save someone else from a charge
    iii. having a weapon

f. Doctor
    i. to not get put on medicine
    ii. to not get surgery
    iii. if your parents are in the room and you don't want them knowing something
    iv. downgrade the severity of your illness/condition/abilities
    v. placebo effect

BIBLIOGRAPHY

[1] Benigni, Roberto, and Vicenzo Cerami. Life Is Beautiful. Miramax, 1999.

[2] Bernstein, Rebecca. “The Mind and Mental Health: How Stress Affects the Brain.” Touro University WorldWide, Touro University WorldWide, 4 Jan. 2019, www.tuw.edu/health/how-stress-affects-the-brain/.

[3] Byrne, C. C., & Kurland, J. A. (2001). Self-deception in an evolutionary game. Journal of Theoretical Biology, 212(4), 457-480.

[4] Carson, Thomas L. Lying and deception: Theory and practice. Oxford University Press, 2010.

[5] Colman, Andrew M., and Eva M. Krockow. "Game Theory and Psychology." Oxford Bibliographies. 27 June 2017. Oxford University Press. 22 Mar. 2019.

[6] Doyen, L. (2017, September). The Multiple Dimensions of Mean-Payoff Games. In International Workshop on Reachability Problems (pp. 1-8). Springer, Cham.

[7] Evans, Gareth. “How Credible Are Lie Detector Tests?” BBC News, BBC, 4 Oct. 2018, www.bbc.com/news/world-us-canada-45736631. poly

[8] Feess, E. & Kerzenmacher, F. (2017). Lying opportunities and incentives to lie: Reference dependence versus reputation.

[9] Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. Journal of the European Economic Association, 11(3), 525-547.

[10] Fishbein, Martin. Progress in Social Psychology. Vol. 1, PSYCHOLOGY Press, 2017.

[11] Hass, Daniel. “Where in the Brain Does Deception Lie?” Lions Talk Science, Penn State College of Medicine, 28 May 2015, lions-talk-science.org/2015/05/28/where-in-the-brain-does-deception-lie/.

[12] Heshmat, Shahram. “What Is Confirmation Bias?” Psychology Today, Sussex Publishers, 23 Apr. 2015, www.psychologytoday.com/us/blog/science-choice/201504/what-is-confirmation-bias.

[13] Ito, Hiromu, et al. "What Is True Halving in the Payoff Matrix of Game Theory?" PloS one 11.8 (2016): e0159670.

[14] Luo, Qiang et al. “The Functional Architecture of the Brain Underlies Strategic Deception in Impression Management” Frontiers in human neuroscience vol. 11 513. 2 Nov. 2017, doi:10.3389/fnhum.2017.00513

[15] Lying opportunities and incentives to lie: Reference dependence versus reputation ∗ Eberhard Feess and Florian Kerzenmacher† November 23, 2017

[16] Lynn, Spencer K and Lisa Feldman Barrett. “"Utilizing" signal detection theory” Psychological science vol. 25,9 (2014): 1663-73.

[17] McEwen, Bruce S and Peter J Gianaros. “Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease” Annals of the New York Academy of Sciences vol. 1186 (2010): 190-222.

[18] Minsky, Marvin Lee. The Society of Mind. Simon & Schuster Paperbacks, 2007.

[19] Murphy, Bill. “How to Tell If Someone Is Lying: 10 Tells and Clues.” Inc.com, Manuseto Ventures., 31 Aug. 2017, www.inc.com/bill-murphy-jr/how-to-tell-if-someone-is-lying-10-tells-and-clues.html.

[20] Park, Alice. (2016). The Fascinating Reason Why Liars Keep on Lying.

[21] Pitts, Leonard. Forward from This Moment. Agate, 2009.

[22] Rajmohan, V and E Mohandas. “The limbic system” Indian journal of psychiatry vol. 49,2 (2007): 132-9.

[23] Sage, J. (2001). The Evolutionary Basis of Self-Deception.1-14.

[24] Shermer, Michael. “The Political Brain.” Scientific American, Scientific American, 1 July 2006, www.scientificamerican.com/article/the-political-brain/.

[25] Stallen, Mirre, and Alan G Sanfey. “The Neuroscience of Social Conformity: Implications for Fundamental and Applied Research.” Frontiers in Neuroscience, Frontiers Media S.A., 28 Sept. 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC4585332/.

[26] Talwalkar, Presh. “How to Beat the Prisoner's Dilemma in the TV Game Show Golden Balls.” Mind Your Decisions, Wordpress, 24 Apr. 2012, mindyourdecisions.com/blog/2012/04/24/how-to-beat-the-prisoners-dilemma-in-the-tv-game-show-golden-balls/. Accessed 27 Mar. 2019.

[27] Temple-Raston, Dina. “Neuroscientist Uses Brain Scan to See Lies Form.” Perelman School of Medicine, University of Pennsylvania, 30 Oct. 2007, www.med.upenn.edu/csa/assets/user-content/documents/NPR_2NeuroscientistUsesBrainScantoSeeLiesForm.pdf.

[28] Trivers, Robert. (1985). Deceit and self-deception. Social Evolution. 395-420.

[29] Trivers, Robert. The Folly of Fools: the Logic of Deceit and Self-Deception in Human Life. Basic Books, a Member of the Perseus Books Group, 2013.

[30] Vann, S D, J P Aggleton, and E A Maguire. "What does the retrosplenial cortex do?" Current neurology and neuroscience reports. Nov. 2009. U.S. National Library of Medicine. 16 Jan. 2019 <https://www.ncbi.nlm.nih.gov/pubmed/198125

[31] Voltaire, et al. Candide. Cideb, 2018.

[32] Wooldridge, Michael J. Reasoning about Rational Agents. MIT Press, 2000

Academic Vita of Elena Cadenas

ecc5222@psu.edu

EDUCATION

Schreyer Honors College at The Pennsylvania State University — **University Park, PA**
Eberly College of Science | Bachelor of Science in Mathematics — *Class of 2019*
Minor in Biology

AWARDS AND SKILLS

Bunton Waller Fellow — *Aug 2015 – May 2019*
- Receive scholarship for academic excellence, which reflects the cultural, racial and ethnic diversity at Penn State

Women in Math Scholarship Recipient — *Aug 2017-May 2018*
Phi Eta Sigma National Honor Society — *Jan 2016- May 2019*
National Society for Leadership and Success — *Aug 2016-May 2019*
Dean's List — *Aug 2015-May 2018*
Skills
- Matlab, HTML, Python and Mathematica programming
- Excel, Tableau, Cytoscape, OriginLab, Gephi, Metascape, Morpheus and Venny 2.1.0 visualization tools
- Amplify DNA samples using polymerase chain reaction, perform gel electrophoresis, and perform ELISA
- Adobe Photoshop and Lightroom
- Proficient with conversational Spanish

RESEARCH

Mathematical Research — *May 2018-May 2019*
- Conduct research to construct a thesis on whether lying is ever truly beneficial
- Conduct a literature review and use multidimensional analysis to make game theory models of various situations in which individuals lie

Mathematical Neuroscience Research Assistant — *Jan 2017-May 2018*
- Research the Combinatorial Threshold-Linear Networks of neural networks
- Use differential equations, graph-theoretic analyses, and Matlab simulations

Participant of the Summer Research Opportunities Program at Purdue University — *Jun 4-Jul 29, 2017*
- Researched and created visualizations on the effect of differentially expressed genes in short and long longevity mouse strains exposed to Diphtheria, Tetanus and acellular Pertussis (DTaP) Vaccine
- Performed ELISA and analyzed the data
- Created multiple visualizations of different styles to demonstrate the results of the data found in the research lab

Cellular Biology Research Assistant — *Sep-Dec 2016*
- Genetically engineered the karst gene through site specific recombination of an engineered intron at an SCE1 site
- Performed primary and secondary screening of larvae to isolate Drosophila melanogaster with the karst gene
- Dissected target larvae to obtain the DNA through cell lysis; performed polymerase chain reaction and gel electrophoresis

WORK EXPERIENCE

Shadowing at Fair Oaks Hospital in the Emergency Department and Family Medicine of Clifton Centreville — *Jun-Aug 2018*
- Interacted with patients, observed lacerations, intubations, and various examinations

Shadowing at Mount Nittany Hospital in Cardiology — *Sep 2017-May 2018*
- Interacted with patients, observed interventional cardiologists, and examined various test results

Shadowing at INOVA Hospital in Radiology — *May-Aug 2016*
- Interacted with patients, watched interventional radiologists, and examined a variety of X-rays

ACTIVITIES

Founder of the Student Success Center Success Coaches — *Sep 2018-May 2019*
- Created the sector of the Success Center that has "Success Coaches" whom act as a support for the student body on all life topics

Remote Area Medical Web Designer and Social Media Executive — *Feb 2017-May 2019*
- Penn State founder. Created website for donations and membership. Update the website, manage all social media, and attend clinics to provide free health care

Schreyer Honors College Scholar Advancement Team — *May 2016-May 2019*
- Ambassador for the Honors College, network with alumni, give tours of Honors housing and campus, and plan events

Scholars Helping Scholars Mentor — *Sep 2016-May 2019*
- Mentor students in Schreyer on various subjects such as homesickness, academics, extracurricular activities, and relationships

Schreyer Honors College Orientation Mentor — *Aug 2016, 2017 and 2018*
- Volunteered and led freshmen transition into life at the Honors College

Global Medical Brigades in Honduras — *Dec 2015-Jan 2016*
- Shadowed medical professionals, conducted triage, gynecology assistance, and dental extraction support
- Learned basic care and provided health care to those in need. Conducted preliminary triage in Spanish with those seeking care

Global Brigades — *Sep 2015 – May 2016*
- Discovered how to use a holistic approach to help underdeveloped countries better their medical care response and infrastructure