THE EFFECTS OF CHEATING ON DECEPTION DETECTION DURING A SOCIAL DILEMMA

by

Kristina L. Swanenburg

B.S., University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

The Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Science in Psychology

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

Kenneth P. Dietrich School of Arts and Sciences

This thesis was presented

by

Kristina L. Swanenburg

It was defended on

May 26, 2011

and approved by

Richard Moreland, PhD, Professor

John Levine, PhD, Professor

Martin Greenberg, PhD, Professor Emeritus

Thesis Director: Richard Moreland, PhD, Professor

THE EFFECTS OF CHEATING ON DECEPTION DETECTION DURING A SOCIAL DILEMMA

Kristina L. Swanenburg, M.S.

University of Pittsburgh, 2011

Research by social psychologists and others consistently finds that people are poor at detecting attempted deception by others. However, Tooby and Cosmides (cognitive psychologists who favor evolutionary analyses of behavior) have argued and shown that humans have evolved a special “cognitive module” for detecting cheaters. Their research suggests that people are good at detecting cheating by group members. These two literatures seem to be at odds with one another. The hypothesis of this research was that when participants are told a lie by a fellow group member whose attempted deception involves cheating on a task that affects their outcomes, they will be good at detecting deception. In this experiment, participants played blackjack in groups using a social dilemma paradigm. Participants’ outcomes were either interdependent or independent with a confederate’s outcomes. It was predicted that participants whose outcomes were interdependent with the confederate would be better at detecting deception by the confederate than those participants whose outcomes were independent from the confederate’s outcomes. Results indicate that when judging other participants’ lies interdependent players were more successful at deception detection than independent players but were not more sensitive to the lies. This effect may be driven by the truth bias, people assume that their interaction partners are truthful which would explain why sensitivity measures (which remove response biases) did not show the hypothesized effect. Independent players were not

more successful or sensitive when judging the confederate's lies. The failure to find the hypothesized effect may be due to methodological factors. Both participants heard may have had their cheating detection modules activated when hearing the instructions for the experiment which implied that cheating could occur. Overall success rates support this idea because they were significantly higher than success rates reached by most deception detection research (50%) which may be indicative that both participants cheating detection modules were active. Results also indicate that as the number of lies told increases overall success decreases but success at detecting lies and sensitivity increase. Thus the more lies that are told the better people are at catching them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Two separate but related research literatures lead to quite different conclusions. The literature on deception detection suggests that people should be poor at detecting cheating exchange partners, whereas the literature on cheating suggests that people are good at detecting such partners. I believe that these literatures differ in their implications because of methodological issues. Research on deception detection often involves participants who share no group context and receive few benefits for catching liars. In contrast, group contexts are made salient for participants in research on cheating, and catching liars may have more benefits for those participants. Social dilemmas offer both group contexts and payoffs that keep participants invested in detecting cheating, similar to the situations in which cheating has been studied. Thus, in the context of a social dilemma, I believe that group members will be good at detecting cheating.

2.0 SOLVING PROBLEMS

How do humans solve the problems that they face? Domain-general or domain-specific solutions can be used. A domain-general solution arises from reasoning processes that can be applied very broadly. These processes are developed through life experiences, and because they are general, can be applied to many kinds of specific problems. For example, one may learn that when faced with large, complex problems, it is best to break them into smaller, simpler parts (see Weick, 1984). This general solution can be applied to a wide variety of problems. Alternatively, people can develop specific solutions for particular problems. These solutions involve cognitive subroutines or "modules" that evolve specifically for that purpose. An example might be the avoidance of vertical drops (Gibson & Walk, 1960). This module is applicable only to the specific problem of vertical drops, where dangerous falls can occur, and has little application to other problems. Cognitive modules develop naturally, without effort, and are often applied unconsciously (Cosmides & Tooby, 2005).

Unconscious application is one of the special strengths of cognitive modules. It helps to resolve the issue of whether everyone has the cognitive ability to consciously apply rules and logic to every problem they encounter. When cognitive modules are used to solve problems, people do not have to engage in complex analyses of those problems. Instead, they can unconsciously and effortlessly come up with the proper solutions. Given all of the complex problems that people face on a daily basis, and the time constraints that they often face in solving those problems, it is likely that people have developed cognitive modules to solve at least some of life's problems.

2.1 Cheating detection module

Cognitive modules probably developed among our ancestors to deal with problems that occurred often in their environment. Because these modules enhanced survival and reproduction, they were more likely to appear repeatedly across generations. Many of the problems that humans have always faced involve social exchange (Cosmides & Tooby, 1989; Cropanzano & Mitchell, 2005; Homans, 1958; Kelley, 1968). Trading food, clothing, or other materials, forming alliances, selecting mates, and cooperating on tasks, all involve exchange processes. Social exchange was likely a recurrent problem for humans - - there is evidence for social exchange among hominids as long as two million years ago (Isaac, 1978). So, there has been plenty of time for cognitive adaptations to social exchange to develop through natural selection (Cosmides & Tooby, 2005). Recent research on von Economo neurons, which are specialized for the processing of information about social interactions (Chen, 2009), suggests that at least one evolutionary adaptation for social exchange exists.

Early humans who developed cognitive modules to assist with social exchange would have benefited from improved reproductive success through access to more and better partners and resources. A major concern in every social exchange is its fairness. Was the exchange fair, or did someone cheat by taking too many resources? If someone were able to detect cheating exchange partners, then he or she could optimize exchange relationships by abandoning cheating partners, swapping cheating partners for partners who did not cheat, or punishing cheating partners to dissuade them from cheating again. In this way, humans in our evolutionary past who could detect cheating by exchange partners probably enjoyed greater reproductive success than did those who could not detect such cheating (Cosmides & Tooby, 2005).

Social exchange is often governed by if-then rules; in order to receive a benefit, one must satisfy a requirement (Cosmides & Tooby, 2005). In a group, for example, one member may wish to exchange fruit for meat possessed by another member; such a person receives the benefit of meat to eat by paying the requirement of providing fruit. For Cosmides and Tooby, cheating involves receiving a benefit without paying the associated cost. For example, someone who received meat without exchanging sufficient fruit for it would be a cheater. Cheating is usually intentional and covert, but it does not have to be.

2.2 Typical paradigm

Because cheating in groups is a recurrent problem, a cognitive module to detect cheating may well have developed. Cosmides (1989) produced evidence for such a module through experiments involving the Wason selection task. In this task, participants search for potential violations of the rule “if P, then Q”. Participants are given four cards to select among, each describing part of the rule on the front side, and a related confirmation or disconfirmation of the rule on the back side. Thus, the front sides of the four cards might read “P”, “not P”, “Q”, and “not Q”. Participants must select which of these cards represents a possible violation of the rule. If a card has “P” written on its front side, then the back side of the card must have “Q” written on it. If a “P” card has “not Q” on its back side, then it is a rule violation. And if a “not Q” card has “P” on its back side, then it is a rule violation too. So, a violation of the rule can only arise on cards that have “P” or “Q” on their front sides. If a card had “not P” written on one side, then it would not matter whether it had “Q” or “not Q” on the other side. Because the rule applies to cases when P is present, the rule “if P, then Q” cannot be violated on this card. In the same way, if a card has “Q” written on one side, then it cannot represent a rule violation either, because whether “P” or “not P” is on the other side of the card, the rule will not be violated. The logical way to detect rule violation in a Wason task is thus to select only the cards that say “P” or “not Q”.

This task is essentially a logic problem, but Cosmides adapted it so that participants had to determine who was (or could be) cheating. This was accomplished by changing the “if P, then Q” rule to “if one receives a benefit, then one must pay a cost”. In Cosmides’ cheating detection paradigm, each card thus indicated whether a benefit was received or a cost was paid. Participants were asked to select the cards in which cheating could have occurred. Participants should have selected only the cards showing that a benefit was received or that a cost was not paid, because these are the two situations in which cheating can occur.

2.3 Typical results

In their research, Cosmides and Tooby compared participants’ accuracy rates (selecting the proper cards on the Wason task) when different frames (descriptions) were applied to the problem. One frame that was often applied involved a descriptive rule using terms that were familiar to participants. Accuracy rates for the Wason selection task when such frames were used were low, ranging from 5% to 30% (Cosmides, 1989). But when a problem was framed in terms of possible cheating during a social exchange (e.g., “if a person borrows a car, then he/she must fill the tank with gas”), accuracy rates were much higher, ranging from 65% to 80%. Cosmides argued that these new frames helped because when people are confronted with a social exchange situation, they can take advantage of their cheating detection module.

The work of Cosmides and Tooby has been criticized; alternate explanations for participant’s performance improvements have been offered. For example, some critics argued that social exchange frames are more familiar to participants and so increased familiarity was confounded with increased knowledge about how to solve the problem. To counter this criticism, Cosmides and Tooby (2005) ran experiments using unfamiliar social contracts, such as, "if a man eats a cassava root, then he must have a tattoo on his face." But even when social contract rules unfamiliar to participants were used, cheating detection accuracy rates were still high, comparable to those for familiar social contracts. Cosmides has addressed other criticisms of her work with Tooby using similar methodological tactics (see Cheng & Holyoak, 1985; Cosmides, 1989).

It should be noted that most of the evidence for the cheating detection module comes from research using the Wason selection task. This lack of methodological variety is worrisome. However, other kinds of studies have also provided evidence that suggests an ability to detect cheaters. For example, research has shown that participants pay more attention to photographs of non-cooperative versus cooperative players in prisoner dilemma games (Vanneste, Verplaetse, Van Hiel, & Braeckman, 2007), and that they can more easily identify photographs of such persons (Verplaetse, Vanneste, & Braeckman, 2007). Mealy, Daood, and Krage (1996) also found evidence of enhanced memory for faces associated with cheating. These studies suggest that people pay more attention to cheaters, which could help them to differentiate cheaters from non-cheaters.

The work by Cosmides and Tooby (2005) suggests that people should be good at catching cheaters. Performance on a simple logic task greatly improves when the task is framed in terms of social contracts in which cheating could occur. However, research on deception detection has led to quite different conclusions.

3.0 DECEPTION DETECTION

Many studies concerning the accuracy of deception detection have been conducted. A recent meta-analysis by Bond and DePaulo (2006) summarized these studies. Most of them (67.4%) involve audiovisual media (e.g., videotapes), in which a college student, discussing various topics, lied half of the time and told the truth half of the time. Participants, usually college students themselves, had to decide when the sender was lying and when he or she was telling the truth. Each videotape typically included a dozen or so messages, most of which (on the average) were less than a minute long. Few of the studies (less than 10%) included actual interaction between the sender and receiver. Bond and DePaulo did not find significant differences in deception detection between studies in which senders and receivers interacted and studies in which they did not. However, their results did show that interaction partners were judged as more truthful by those with whom they interacted than by those with whom they did not interact. The authors believe that this occurs because people are reluctant to view those with whom they interact as liars.

The Bond and DePaulo (2006) meta-analysis found that the accuracy of participants at detecting deception was 54%, on the average. Considering that most people believe they are good at detecting deception, an accuracy rate that is little better than chance (50%) seems surprising. This inability to detect deception is not just an issue for college students. Ekman and O'Sullivan (1991) found that even people in occupations that rely on accurate deception detection, such as judges and police officers, do little better than chance at detecting lies. The best performing group in their study, Secret Service agents, only managed a 64% accuracy rate.

One explanation for poor deception detection is that people often pay attention to the wrong cues. For example, people commonly believe that liars are fidgety and feel guilty about the lies they are telling. These misconceptions lead people to focus on the wrong cues. In fact, Vrij and Mann (2001) found the worst levels of deception detection among people who relied on such cues as fidgeting and gaze aversion. Some of the more accurate deception clues (e.g., the use of gestures to demonstrate what is being said) are not connected to a liar's feelings of guilt. Thus, some of the best cues for detecting deception are actually ignored by people, making them poor lie detectors.

Another reason why people are poor at detecting lies is the truth bias--the tendency for people to be more accurate at determining that a truth is true than they are at determining that a lie is false (Zuckerman, Koestner, Colella, & Alton, 1984). In other words, people tend to assume that what they are told is true. Research on the truth bias suggests that the best predictor of accuracy in detecting deception is whether messages are actually true or false (Levine, Park, McCornack, 1999).

3.1 Methodological problems in deception detection research

Deception detection research, like research on cheating, is also plagued by a lack of methodological variety. Many of the cues that people might use in real-life situations to detect deception are not available in this research, because there is no direct interaction between senders and receivers. Also, receivers and senders rarely know one another. This could limit the accuracy of the receivers' judgments. As their familiarity with someone increases, people may well learn how to differentiate that person's lies from truths. Bond and DePaulo (2006) also noted that in real-world situations, there is usually more time than just a few minutes to discover an untruth.

More importantly, the lies that senders tell are usually irrelevant to the receivers' outcomes in this research. Some of the more common messages involve senders' political views, work history, or family details, all things that do not affect the receivers. And the motivation of receivers to detect deception has been mostly ignored in this research, whereas the motivation of senders to successfully deceive receivers has received close attention. In at least 40% of the studies examined by Bond & DePaulo (2006), the motivation of senders was acknowledged as a factor. Methods for increasing senders' motivation (see Ekman & O'Sullivan, 1991; Frank and Ekman, 1997) have included both positive and negative reinforcement (monetary gain and pain avoidance, respectively) of senders for successful deception, and persuading senders that successful deception is linked to career success. Research interest in the sender's motivation was stimulated by Ekman's theories about how emotionality affects the cues exhibited by liars. Ekman argued that high stakes increase the chances that deception clues will be "leaked" by liars (Frank & Ekman, 1997). Indeed, much research has shown that the more motivated senders are, the worse they are at deceiving receivers (DePaulo, Kirkendol, Tang, & O'Brien, 1988; Forrest & Feldman, 2000).

4.0 WHY THE DISCREPANCIES BETWEEN THE LITERATURES?

Why does research on cheating suggest that people can detect cheaters, whereas research on deception detection suggests that people are poor at detecting lies? There are many possible reasons for this discrepancy. Two of the most important reasons may be group context and personal relevance. The group context is apparent in the cheating literature (the use of social contracts in the Wason task), whereas groups are largely ignored in the deception detection literature. A second discrepancy between the two literatures is personal relevance. As mentioned earlier, few of the lies/truths examined in deception detection research are personally relevant to the receiver; the messages are about the sender more often than about the receiver, and they rarely have any significance to the receiver. Research on hedonic relevance (Jones & Davis, 1965) suggests that personal relevance increases the likelihood that someone will make attributions, so it could also be important when detecting deception. For example, when a statement is more relevant to someone, he or she may be more likely to make attributions about the sender (like the person is dishonest). Cosmides and Tooby (2005) do not use paradigms that directly involve personal relevance, but people often have a general investment in others' adherence to social contracts. If a social contract is broken by one person, then it could presumably be broken by others, harming everyone. So, participants may be more concerned with the breaking of social contracts in cheating detection research than participants in deception detection research are with detecting irrelevant lies.

5.0 SOCIAL DILEMMAS

Social dilemmas involve conflicts between individual and collective goals. In social dilemmas, individuals receive benefits for behaving in selfish ways, but if everyone behaves selfishly, then they all lose resources (Dawes & Messick, 2000). Generally, two types of social dilemmas are studied: public goods dilemmas and common pool dilemmas. Public television is a good example of a public goods dilemma. Public television is a good to which everyone has access, a good that is supported by voluntary donations. Not everyone has to donate, but if too few people donate, then the good can no longer be supported and everyone loses access to it. Common pool dilemmas are often compared to limited natural resources, such as timber or fish. There is a naturally replenishing resource to which everyone has access. Individuals can take as much of that resource as they want, but if too much of the resource is taken, then the resource can no longer replenish itself and everyone loses access to it.

Public goods dilemmas may evoke the cheating detection module proposed by Cosmides and Tooby more readily than do common pool dilemmas. Public goods dilemmas clearly involve taking a benefit without paying a cost (e.g., watching public television without having donated). Common pool dilemmas can also involve cheating, but the rule of taking a benefit without paying a cost is less relevant to such dilemmas. So, a public goods social dilemma may be better than a common pool dilemma for studying cheating (as defined by Cosmides and Tooby) and its detection.

6.0 METHODOLOGY

I created a public goods dilemma in the laboratory to study cheating occurrence and deception detection in groups. In my experiment, personal relevance (of the cheating and deception) was manipulated as an independent variable. If participants have a reason to be concerned about lies being told, because cheating could occur and might affect their personal outcomes, then their cheating detection modules are more likely to be activated. As a result, they should be able to detect lies more accurately. A second independent variable in my experiment was the number of lies that were told. A detection module is more likely to be activated as more lies are told. Finally, there may also be an interaction effect involving the personal relevance of lies and the number of lies that are told. The positive impact on deception detection of the number of lies that are told by the confederate should be stronger among interdependent players than among independent ones. The activation of the cheating detection module will make interdependent players more likely to detect lies as more lies are told.

My experiment followed a (2) x 3 design. The first independent variable, which varied within groups, contrasted group members whose outcomes in the social dilemma were independent with those whose outcomes were interdependent. This manipulated the personal relevance of cheating. The second independent variable, which varied between groups, involved the number of lies that were told (zero, three, or six) by a confederate posing as a group member.

Three-person, same-sex groups were studied, each containing two real participants and the confederate. Because there are no clear differences in deception detection accuracy between males and females (DePaulo, Epstein, & Wyer, 1993), all of the groups were composed entirely of females. Restricting the study to one gender simplified the data analyses and lowered my research costs.

Participants were asked to report on their behavior during the social dilemma, including possible cheating. It was unclear to me, before the experiment began, how willing participants would be to cheat, or whether they would try to deceive others about their cheating. That is why

I added a confederate to each group, someone who was instructed to cheat and lie according to a pre-determined schedule.

After they arrived at the laboratory, group members first read and signed informed consent forms. These forms specified only that the study was about behavioral judgments and decision making. Participants were then seated around a single table. Each person was given a nametag with a letter (indicating her role) on it, so that she could be readily identified by the other participants. The potential cash that participants could win was visible on a nearby table, to strengthen participants' motivation to attain good personal outcomes during the experiment. The general rules of blackjack were then explained to participants and a sample blackjack hand was played with the group. Blackjack is a familiar and enjoyable card game, which helped to maintain participants' interest and motivation. Once participants understood the general rules of blackjack, the specific betting rules for the experiment were explained, and a practice hand was played under those rules.

During each round of the game, any group member could place a bet against the dealer (the experimenter). If someone beat the dealer, then the money that was collectively bet by the group as a whole was doubled and given to the players. If a person did not bet, then she received all of that money. If more than one person did not bet, then the money was divided evenly among the non-betters. And if no one bet, then there was no money to divide, so no one received any money whether the players won (because their hands were better than the dealer's hand) or not. If everyone bet, then the money was divided evenly among them. Thus, there was no additional benefit to more than one player winning against the dealer. If no one won, then the dealer kept all of the money that was bet.

These betting rules created a kind of public goods social dilemma. An individual player did best by not betting, because that player could then keep more of her money and still collect money from the other group members (if other bets were made). However, if all group members acted in this selfish way, then no one collected any winnings. This situation also fits the cheating scenario of taking a benefit without paying a cost -- cheating here would involve collecting winnings without betting any money.

The betting rules that a participant followed depended on her role in the experiment. One participant in each group (Player C) was randomly selected to be independent; the other participant (Player B) and the confederate (Player A) were interdependent. So, although everyone in the group played blackjack, Player C was unaffected by the other players' bets. In other words, Player C could bet on each round (just like the other players), but could not lose money to other players who chose to cheat, or win other players' money by cheating. Participants were aware of one another's roles and betting rules. For a summary of the betting scenarios that could occur, and the corresponding amounts of money that would be won by different group members, see Table 1.

Table 1. Possible bets (and winnings) for the members of a group

Bets			Winnings		
Member A	Member B	Member C	Member A	Member B	Member C
Yes	Yes	Yes	$0.50	$0.50	$0.50
No	Yes	Yes	$1.00	------	$0.50
Yes	No	Yes	------	$1.00	$0.50
No	No	Yes	------	------	$0.50
No	No	No	------	------	------
Yes	Yes	No	$0.50	$0.50	------
No	Yes	No	$1.00	------	------
Yes	No	No	------	$1.00	------

To help participants understand the betting rules, they were asked to fill out a worksheet containing several questions about how much money each player would win given different betting scenarios and blackjack hands. After reviewing the worksheet for ten minutes, few participants had trouble understanding the rules. Participants who did have trouble were given repeated instructions on the betting rules until they understood them. Ten rounds were played. On each round, participants could choose to bet or not bet. Every bet was worth 25 cents. All bets were placed privately, so other group members were unaware of whether a person was betting. This was accomplished by giving an index card to each participant for every round. The index cards were marked with a symbol to indicate the round in which they were used. Symbols were used (instead of round numbers) to weaken players' concerns about their betting behavior being monitored during each round. The participant wrote on each card whether she wanted to bet and then placed the card in a box with opaque sides. In this way, the experimenter could determine, after the experiment, how the participant actually bet.

After each round, the participants were asked to say aloud, in a full sentence, whether they had bet on that round. Participants were told that they did not have to be truthful when describing their betting behavior. After each group member spoke, the other players indicated on a questionnaire whether they thought she had actually bet or not. Participants were also asked to rate how confident they were about these judgments (people may be less accurate at detecting lies when they are less certain). Finally, to ensure that participants paid attention to what other group members said, they were also asked whether each person said that she had bet (or not) on each round.

After all of the rounds were done, participants completed a group identification scale (Hinkle, Taylor, & Fox-Cardamone, & Crook, 1989). The more that a participant identifies with a group, the less willing she may be to judge other group members as liars. This scale contains nine items that reflect the emotional, cognitive, and behavioral aspects of social identity. Participants rated each item on a five-point scale that ranged from 1 (strongly disagree) to 5 (strongly agree). A manipulation check on personal relevance was also administered, after the group identification scale. On a scale from 1 (not at all) to 5 (completely), participants were asked, "How much were your outcomes affected by other members of your group?" Finally, one question ("yes" or "no") was also asked about any prior gambling experience that participants may have had. Those with gambling experience may be more likely to lie.

After all of these measures were completed, participants were probed for suspicions about the experiment. The confederate was then excused from the room and the experimenter explained the role of the confederate in the experiment and why it was necessary to use a confederate. During the rest of the debriefing, care was taken to ensure that participants realized that any deception and cheating by others was for the economic gain of those persons and did not represent personal insults or attacks on the participants themselves. Also, the experimenter made clear that choosing to "cheat" in an artificial social dilemma does not necessarily reflect someone's morality outside of the laboratory.

As noted earlier, the experiment utilized a confederate. A female confederate was chosen, someone who resembled the kinds of students who participated in the experiment. She was trained during several one-on-one sessions with the experimenter. These sessions involved a review of the methodology and some practice at deception. The confederate was told to display consistent behavior in every group during the experiment. Several practice sessions with participants were conducted to complete the confederate's training. These sessions included every level of the number of lies variable (zero, three, and six). The data from these sessions were not included in later analyses.

During each session of the actual experiment, the confederate only bet twice, on the third and ninth rounds. When asked to report her betting behavior, she sometimes tried to deceive group members. The confederate lied a different number of times, depending on condition. In the zero lies condition, the confederate was always honest about her betting behavior. In the three lies condition, the confederate was dishonest about her betting behavior on the second, fifth, and eighth rounds (she failed to bet eight times, but said that she did not bet five times). In the six lies condition, the confederate was dishonest about her betting behavior on the second, fourth, fifth, sixth, eighth, and tenth rounds (she failed to bet eight times, but said that she did not bet twice). This schedule reflected an effort to "spread out" when the confederate's lying occurred. Every lie told by the confederate involved a round in which she cheated – she never lied about not cheating. Future research could investigate the detection of that kind of deception, but it was not the focus of my experiment.

The use of a confederate was not ideal, because the lies that she told may have been in a sense more "artificial." A confederate's deception cues may thus differ from those of people who cheat and lie of their own accord. However, a confederate was necessary because participants might assume that honesty is "proper behavior" in a laboratory experiment, leading them to lie rarely or never. Another important issue was the willingness of participants to cheat in this paradigm. Other researchers have found that participants in prisoner's dilemmas (a type of social dilemma) only cooperate about half of the time (Komorita & Parks, 1996). Furthermore, a meta-analysis of research on several types of social dilemmas found the mean cooperation rate to only be 47% (Sally, 1995). Thus, cheating is fairly common in the work of other researchers, but that did not guarantee that it would occur often in my own research.

Lying about cheating creates an opportunity for deception detection. With the use of the cheating detection module, participants should be able to determine when someone is lying to them. Participants whose outcomes were actually affected by the cheating had more reason to try to accurately detect deception. This reasoning led to the following hypothesis:

Hypothesis 1: Participants will show greater accuracy at detecting the confederate's deception when they are interdependent with the confederate than when they are not.

As described earlier, people should more easily detect lies when more lies are told. This effect should be even stronger for those who are interdependent with the confederate than for those who are not interdependent with her. This reasoning led to the following hypotheses:

Hypothesis 2: Participants will show greater accuracy at detecting the confederate's deception when she tells more lies.

Hypothesis 3: The positive impact on deception detection of the number of lies that are told by the confederate will be stronger among interdependent players than among independent ones.

Finally, an interesting research question involves the possible contagion of cheating in the groups:

Research Question 1: Would cheating and lying by the confederate encourage cheating and lying by other group members? And when a confederate lies, would that improve or worsen the accuracy with which group members judge one another's lies?

7.0 RESULTS

A total of 66 sessions were run. Five of these were pilot sessions that helped the confederate to practice her script and become comfortable with the experimental procedures. One group's data were later discarded after a participant expressed suspicion about the confederate. This left a total of 60 groups: 20 in the zero lies condition, 20 in the three lies condition, and 20 in the six lies condition. There were two real participants in each group, so a total of 120 participants were in the experiment.

7.1 Participant awareness

The question used to ensure that participants were paying attention to other's responses indicated that participants were indeed listening to each other. The average number of mistakes made – when participants inaccurately recorded what other participants said (over ten rounds and twenty possible responses), was only 0.42 (SD = 1.02) with a range of 0 to 8. Participants thus made less than one mistake, on average. Eighty-two participants made no mistakes at all. So, participants indeed seemed to be paying attention to what other group members were saying.

7.2 Manipulation checks

As noted earlier, participants were asked to rate how much their outcomes were affected by other group members. Recall that participants in the Player B role were interdependent with the confederate (Player A), whereas those in the Player C role were not. On average, participants in the Player B role scored a 3.27 (SD = .97) on the scale (the midpoint was 3.0), whereas those in the Player C role scored a 2.52 (SD = 1.08). Both means were significantly different ($p < .05$) from 3.00. A (2) x 3 mixed ANOVA (using the group as the unit of analysis) showed a significant effect for role, $F(1, 57) = 21.34$, $p < .01$, but no significant effect for the number of lies told by the confederate, $F(2, 57) = 1.13$, $p > .05$, or for the interaction between role and the number of lies that the confederate told, $F(2, 57) = .98$, $p > .05$.[1] These results indicate that the interdependence manipulation was successful.

There was no need to check the effectiveness with which the number of lies told by the confederate was manipulated – the confederate followed her instructions carefully, failing to bet and lying about that behavior the correct number of times, and during the correct rounds, in every group. However, the participants themselves sometimes lied. Forty-four participants, or about a third of the sample, told no lies at all. But the remaining participants did lie, some as many as eight times during the ten rounds of the experiment. The mean number of lies told by participants was 1.67, with a standard deviation of 1.85. Did the number of lies told by participants vary across the cells of the design? To explore this issue, a (2) x 3 ANOVA was performed. The results showed that the number of lies told by participants did not vary by their role, $F(1, 57) = 2.76$, $p > .05$, or by the number of lies that the confederate told, $F(2, 57) = 1.18$, $p > .05$. The interaction between role and the number of lies told by the confederate was not significant either, $F(2, 57) = .10$, $p > .05$. All of this suggests that only the confederate's lies, and not the lies told by the participants themselves, had to be included in subsequent analyses. It also provided a partial answer to my research question – lying by the confederate did not seem to encourage lying by the participants. [2]

7.3 Confederate behaviors

Did the confederate indeed behave consistently throughout the experiment, as she was encouraged to do? To investigate this issue, two coders were recruited to evaluate videotapes made of three groups (one from each condition). These coders were psychology graduate students, blind to my research hypotheses and to the condition that each videotape represented. While watching each videotape, they independently rated the confederate on a five-point scale (1

= low, 3 = moderate, 5 = high) for behavioral consistency, vocal consistency, and believability on each round. Behavioral consistency involved the gestures or body language that the confederate used while speaking. Vocal consistency referred to her tone, speed of speech, and general vocal quality. The coders' scores, which were quite similar (the ICC for the judges, across all three behavioral categories was $\alpha = .64$), were then averaged together to produce a single score for each aspect of the confederate's behavior. These scores (on a five-point scale) were high for behavioral consistency ($M = 4.97$, SD = .19), vocal consistency ($M = 4.93$, SD = .26), and believability ($M = 4.89$, SD = .31). All of the means were significantly greater than 3.00, the midpoint of the rating scales, suggesting that the confederate indeed performed her role successfully.

7.4 Group identification and gambling experience

Coefficient alpha for the items on the group identification scale was .78, indicating an adequate level of reliability. The mean rating across items was thus calculated to create a single group identification score for each participant. These scores ranged from 2.11 to 4.67, with a mean of 3.40 and a standard deviation of 0.48. This mean was significantly higher than 3.0, the midpoint of the scale, $t(119) = 9.31$, $p < .01$. This indicated that group members identified strongly with their groups. However, the results from a (2) x 3 ANOVA indicated that group identification was not significantly related to a player's role, $F(1, 57) = .40$, $p > .05$, the number of lies told by the confederate, $F(2, 57) = .64$, $p > .05$, or the interaction between those two factors, $F(2, 57) = .64$, $p > .05$. This suggested that group identification did not need to be included as a factor in subsequent analyses.

Recall that participants were also asked whether they had previous gambling experience. Only 27 of the 120 participants had such experiences. These participants were given a score of "1" on a prior experience variable; everyone else was given a score of "0" on that variable. A (2) x 3 analysis of variance was then performed, using these scores as the dependent variable. The results showed that gambling scores were not related to a player's role, $F(1, 57) = 0.05$, $p > .05$, the number of lies that the confederate told, $F(2, 57) = 0.42$, $p > .05$, or the interaction between those two variables, $F(2, 57) = .32$, $p > .05$. Thus, previous gambling experience was not included as a factor in subsequent analyses either.

7.5 Winnings

There were two possible ways to measure a participant's winnings. One was to count how many rounds (out of ten) the person won. The other was to calculate the total amount of money the person won. On average, each player was successful, winning 6.48 rounds, with a standard deviation of 1.54. A repeated measures analysis of variance, with the number of lies told by the confederate as the only independent variable, was performed using the number of rounds won as the dependent variable. [The role variable was not included in this analysis, because both of the participants in each group always won or lost together]. The results showed that the number of lies that a confederate told had no effect on the number of rounds that players won, $F(2, 57) = 2.68$, $p > .05$. However, this way of measuring winning is less precise than evaluating the amount of money that each player won, because the latter measure can vary across players within a group. An additional analysis was thus performed, using the amount of money that a player won as the dependent measure.

The average amount of money that a player won was \$2.49, with a standard deviation of \$0.93. A (2) x 3 ANOVA using this measure showed a significant effect for player role, $F(1, 57) = 127.58$, $p < .01$. Participants in the Player B (interdependent) role won only \$1.85 on average, whereas those in the Player C (independent) role won \$3.13 on average. Participants in the Player B role probably won less money because the confederate's cheating affected their outcomes. This finding suggests that participants in that role were not very good at detecting deception by the confederate. Indeed, the number of lies that the confederate told significantly affected the amount of money that a player won, $F(2, 57) = 5.87$, $p < .01$. On average, players won \$2.76 (*S.D.* = \$0.11) when the confederate told zero lies, \$2.46 (*S.D.* = \$0.11) when the confederate told three lies, and \$2.25 (*S.D.* = \$0.11) when the confederate told six lies. A Scheffe's post-hoc analysis showed that players won significantly ($p < .01$) more money when the confederate told zero lies than when she told six lies; none of the other differences in winnings was significant. Finally, the interaction between a player's role and the number of lies that the confederate told did not significantly affect the amount of money that a player won, $F(2, 57) = .10$, $p > .05$. The amount of money won was unrelated to the accuracy with which players detected lies, and so it was not included in any later analyses.

7.6 Deception detection

There are several ways to measure the accuracy of deception detection. One involves the calculation of a simple "success" variable, where success is defined as the number of trials (out of 10) in which someone correctly classified a person's claim about betting as true (when the claim was true) or false (when the claim was a lie). Because this variable is a percentage, it was transformed using the arcsine transformation (to normalize its distribution) before being used in any analyses. The sample mean success score was 56.42%, suggesting that participants were detecting deception at slightly better than the chance rate of 50.00%, $t(119) = 3.22$, $p < .01$. This is different from most deception detection research which finds deception detection at chance levels.

A (2) 3 ANOVA was then carried out on success scores. The independent variables were Player Role (B or C), which was a within-groups variable, and the Number of Lies told by the confederate (0, 3, or 6), which was a between-groups variable. In this analysis, Player Role was not a significant factor, $F(1, 57) = 1.79$, $p > .05$. However, the Number of Lies told by the confederate was a significant factor, $F(2, 57) = 37.51$, $p < .01$. The mean success score was 1.14 (*S.D.* = .04) when the confederate told no lies; 0.76 (*S.D.* = .04) when the confederate told three lies; and 0.72 (*S.D.* = .04) when the confederate told six lies. These scores are equivalent to success rates of 82.56%, 47.46%, and 43.48% respectively. Thus, participants generally had more trouble determining whether the confederate was lying when she told more lies. Scheffe's posthoc tests showed that success scores were significantly ($p < .01$) higher when the confederate told zero lies than when she told three lies, and significantly higher when she told zero lies than when she told six lies. There was no significant difference, however, between success scores when the confederate told three versus six lies. Finally, the interaction between Player Role and Number of Lies told by the confederate was not a significant factor, $F(2, 57) = .51$, $p > .05$.

One problem with the analysis just described is that the success scores confounded two kinds of success - - success at determining when someone was telling the truth and success at determining when someone was lying. To distinguish between these two kinds of success, a further analysis was performed, a (2) x (2) x 3 ANOVA. It was identical to the analysis just described, except that Type of Success was added as a new independent variable. This was a within-groups variable, just like Player Role. Each player thus had two success scores, one calculated as the percentage of times when the confederate told the truth and the participant realized that she was truthful, and the other calculated as the percentage of times when the confederate lied and the player realized that she lied. Because they were also proportions, these new scores again underwent an arcsine transformation before analyzed.

The new ANOVA showed that Player Role did not significantly change success rates, $F(1, 55) = .73$, $p > .05$. However, success rates did vary significantly with Type of Success, $F(1, 55) = 311.04$, $p < .01$. Participants were significantly better at detecting truths ($M = 1.08$, $SD = .26$; 77.79%) than they were at detecting lies ($M = .21$, $SD = .27$; 4.35%). And Number of Lies was again a significant factor, $F(2, 55) = 1349.98$, $p < .01$. Scheffe's posthoc analyses showed that success rates were significantly higher when the confederate told six lies ($M = .79$, $SD = .03$; 50.46%) than when she told three lies ($M = .58$, $SD = .03$; 30.03%) or zero lies ($M = .57$, $SD = .03$; 29.12%). The differences in success rates when the confederate told zero and three lies were not significant. The two-way interaction between Player Role and Type of Success was not significant, $F(1, 55) = .06$, $p > .05$. The two-way interaction between Player Role and the Number of Lies was not significant either, $F(2, 55) = .18$, $p > .05$. However, the two-way interaction between Type of Success and Number of Lies *was* significant, $F(2, 55) = 7.57$, $p < .01$ (See Figure 1). As predicted, participants were better at detecting lies as Number of Lies increased. But the detection rate of truths showed a different pattern. When the confederate told zero lies, success at detecting truths was relatively high. But success decreased when the confederate told three lies. When the confederate told six lies, success rates regained their former level. Simple effects analyses showed that the success rate at detecting truths was significantly better than the success rate at detecting lies at all three levels of confederate lies (zero, $F(1, 55) = 170.74$, $p < .01$; three, $F(1, 55) = 62.60$, $p < .01$; and six, $F(1, 55) = 90.49$, $p < .01$). The three-way interaction between Player Role, Type of Success, and Number of Lies was not significant, $F(2, 55) = 1.12$, $p > .05$.

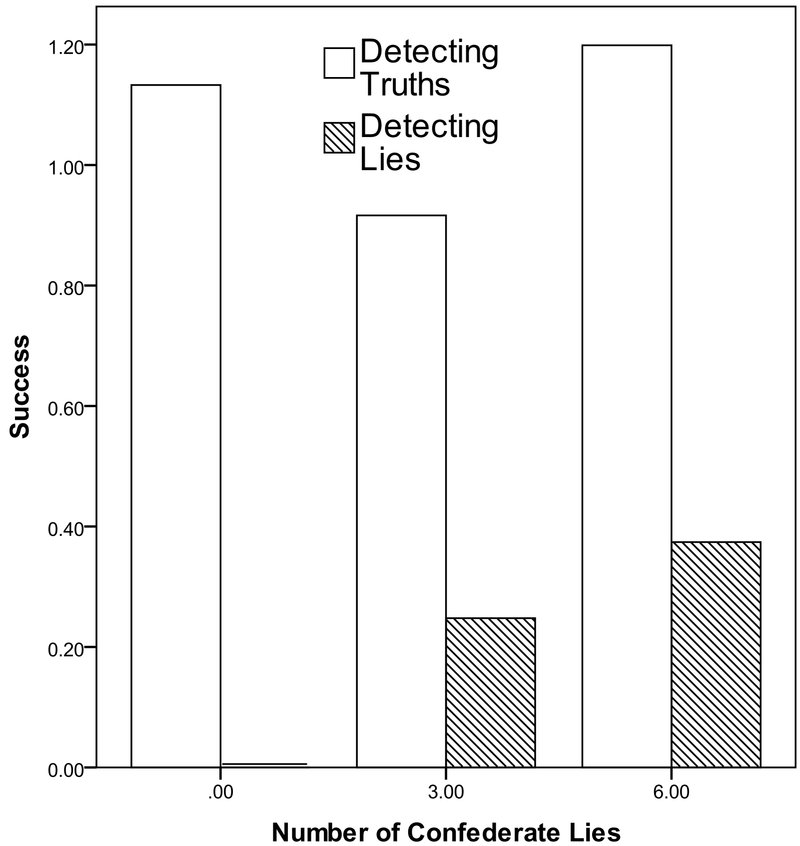

Figure 1. The interaction between Type of Success and the Number of Lies told by the confederate on success (arcsine transformed) at detecting truths and lies.

Another problem with the success variable was that it did not allow for the separation of two factors that might affect detection accuracy, namely sensitivity and response bias. In my research context, sensitivity refers to a person's ability to somehow know a lie when it is spoken, whereas response bias refers to a person's tendency to systematically under- or over-estimate how often lies are told. Some people, for instance, may assume that no one lies, whereas others assume that everyone lies (Rotter, 1971).

Signal detection analysis is a statistical method that can independently assess the levels of sensitivity and response bias within a data set. To conduct a signal detection analysis, one must calculate the "hit" rate, the "miss" rate, the "false alarm" rate, and the "correct rejection" rate. For my research, a hit would be a correctly detected lie, a miss would be a missed lie (a lie viewed mistakenly as the truth), a false alarm would be an incorrectly detected lie (the truth viewed as a lie), and a correct rejection would be a correctly accepted truth (a truth viewed mistakenly as the truth). One can use the hit rate and the false alarm rate to calculate D-prime (d'), a statistic that measures sensitivity. Because the hit rate and false alarm rate may have different distributions, they must both be transformed to z-scores. D-prime is then calculated by subtracting the z-score for the false alarm rate from the z-score for the hit rate. The higher (more positive) the d' is, the more sensitive participants are to the signal (lies). If the hit rate is equal to the false alarm rate, then d' is zero.

Response bias was calculated using some of the same information used to calculate d'. The z-score for hits and the z-score for false alarms were first added together. This sum was divided by two and the result was multiplied by negative one. Response bias values indicate the tendency of the participant to judge statements as lies or truths. A response bias of zero indicates that the participant had no tendency in either direction. Positive values of response bias indicate that a participant was likely to assume that what she heard was the truth, and negative values indicate that a participant was likely to assume that what she heard was a lie. One drawback of signal detection analysis for my research was that lying had to occur in order for d' to be calculated at all. Thus, the zero lies condition could not be included in the analyses involving confederate lies.

In the first analysis, a (2) x 2 ANOVA of sensitivity scores was performed, with Player Role (Player B or C) as a within-groups variable and Number of Lies told by the confederate (3 or 6) as a between-groups variable. The average sensitivity score was .02, which was not significantly different from zero, $t(39) = .1$, $p > .05$. There was no significant effect of Player Role in this analysis, $F(1, 38) = .29$, $p > .05$, but the Number of Lies told by the confederate was a significant factor, $F(1, 38) = 14.81$, $p < .01$. As predicted, there was greater sensitivity in the six lies condition ($M = .50$, $SD = .18$) than in the three lies condition ($M = -.47$, $SD = .18$). The interaction between Player Role and Number of Lies told by the confederate was not significant, $F(1, 38) = .25$, $p > .05$.

In the second analysis, the same (2) x 2 ANOVA was performed, but with response bias as the dependent variable. No significant effects were found for Player Role, $F(1, 38) = .48$, $p > .05$, the Number of Lies told by the confederate, $F(1, 38) = .68$, $p > .05$, or their interaction, $F(1, 38) = 1.96$, $p > .05$.

7.7 Confidence

Recall that when participants decided whether the confederate was lying or not, they also rated their confidence in those judgments on a 1 to 5 scale. The average confidence rating (across all 10 trials) was computed and found to be 3.36 ($SD = 0.69$), significantly higher than 3.0, the midpoint of the scale, $t(119) = 5.75$, $p < .01$. This rating suggests that participants had confidence in their judgments. Correlational analyses showed that these average confidence ratings were not correlated with success, $r = -.02$, $p > .05$. Participants' success at detecting lies thus had no relationship with their levels of confidence. However, the correlations between confidence and the two Types of detection success were significant. There was a positive correlation between confidence and success at determining when someone was telling the truth, $r = .29$, $p < .01$, and a negative correlation between confidence and success at determining when someone was lying, $r = -.34$, $p < .01$. The truth bias could account for these correlations. The truth bias made participants more likely to think that their groupmates were telling the truth. Participants who judged their groupmates as truthful were more confident because of this supposed honesty. And their judgments were more likely correct when others were truthful and incorrect when others lied. Confidence scores and sensitivity scores were not significantly correlated, $r = -.11$, $p > .05$. But confidence scores were significantly correlated with response bias scores, $r = .45$, $p < .01$. Participants with higher levels of bias (indicating that they viewed more statements as truths) were more confident. Participants were thus more confident insofar as they assumed that they were being told the truth.

A (2) x 3 ANOVA was performed on the participants' confidence scores, using Player Role as a within-groups factor, and the Number of Lies told by the confederate as a between-groups factor. The results showed that Player Role was not a significant factor, $F(1, 57) = .02$, $p > .05$, nor was Number of Lies, $F(2, 57) = 1.49$, $p > .05$, or the interaction between these factors, $F(2, 57) = .76$, $p > .05$.

7.8 Participant lies

The previous analyses all involved only confederate lies. However, as noted earlier, participants were allowed to (and often did) lie. In each group, an average of 3.33 (*SD* = 2.60) lies were told by participants (with a range of 0 to 11). When the number of lies told by participants in each group was added as a covariate to the (2) x 3 ANOVA on success, the covariate was not significant, $F(1, 56) = .78$, $p > .05$. The overall findings of this ANOVA were also unchanged with the addition of the covariate. When the number of lies told within each group was added as a covariate to the ANOVA of sensitivity scores, the covariate was again not significant, $F(1, 37) = .40$, $p > .05$. The findings of this ANOVA were also unchanged with the addition of the covariate. The findings of the ANOVA of response bias scores were also the same after the addition of the covariate. However, the covariate *was* significant in this case, $F(1, 37) = 11.79$, $p < .01$. As the number of lies told in the group increased, Player B showed less response bias and Player C showed more response bias (although this interaction was not significant).

To investigate participant lies further, groups were divided into three sets, based on how many lies their members told. The first set contained groups (N = 18) in which zero or one lie was told. The second set contained groups (N = 22) in which two, three, or four lies were told. Finally, the third set contained groups (N = 20) in which five or more lies were told. It should be noted that participant lies could vary more than the confederate's lies. The confederate only lied by saying that she bet when she did not bet. However, participants could lie by saying that they bet when they did not bet or by saying that they did not bet when they actually did bet.

In terms of overall ability to detect deception, participant lies produced mixed results. The average success score was 1.07; 76.77% (*SD* = .29), which is significantly higher than 50%,

$t(119) = 10.58$, $p < .01$. But the average sensitivity score was -.12 ($SD = 1.55$), which is not significantly different from zero, $t(119) = -.86$, $p > .05$.

A (2) x 3 ANOVA on success scores was performed, with Player Role as the within-groups variable and Number of Participant Lies (low, medium, or high) as the between-groups variable. Player Role was significant, $F(1, 57) = 6.67$, $p < .05$. Participants in the Player B role ($M = 1.14$, $SD = .04$) had higher success scores than did participants in the Player C role ($M = 1.01$, $SD = .03$). These means reflect success rates of 82.56% and 71.71% respectively; participants in the Player B role were thus more successful when judging statements as truths or lies than were participants in the Player C role. The Number of Participant Lies told was also a significant factor, $F(2, 57) = 11.09$, $p < .01$. Scheffe's post-hoc tests show that there were significant differences between the low ($M = 1.22$, $SD = .04$; 88.19%) and the medium groups ($M = 1.06$, $SD = .04$; 76.10%), and between the low and the high groups ($M = .94$, $SD = .04$; 65.22%). Thus, as more lies were told by group members, participants were less successful at deception detection. This finding does not support the second hypothesis. The interaction between Player Role and the Number of Participant Lies was not significant, $F(2, 57) = .27$, $p > .05$.

Next, the same (2) x 3 ANOVA was run on sensitivity scores. Player Role was not significant, $F(1, 57) = .01$, $p > .05$, but Number of Participant Lies was again significant, $F(2, 57) = 11.9$, $p < .01$. Scheffe's post-hoc tests show that there were significant differences between the high ($M = .77$, $SD = .23$) and the low groups ($M = -.73$, $SD = .24$), and between the high and the medium groups ($M = -.44$, $SD = .22$). Thus in groups with lots of participant lying, sensitivity was higher. This finding supports the second hypothesis. The interaction between Player Role and Number of Participant Lies was significant, $F(2, 57) = 3.28$, $p < .05$ (See Figure

2). Simple effects analyses showed that as participant lying increased, there were marginally significant increases in sensitivity scores for Player B, $F(2, 57) = 2.49$, $p < .1$, and significant increases in sensitivity scores for Player C, $F(2, 57) = 11.52$, $p < .01$. Further simple effects analyses showed marginally significant differences between Player B and Player C on sensitivity scores when the Number of Participant Lies was high, $F(1, 57) = 3.94$, $p < .1$. Player C had higher sensitivity scores than Player B under these conditions. Neither player was dependent on the other, so this difference may indicate that Player B focused more on the confederate's statements than on Player C's statements, because Player B was interdependent with the confederate, but not with Player C.

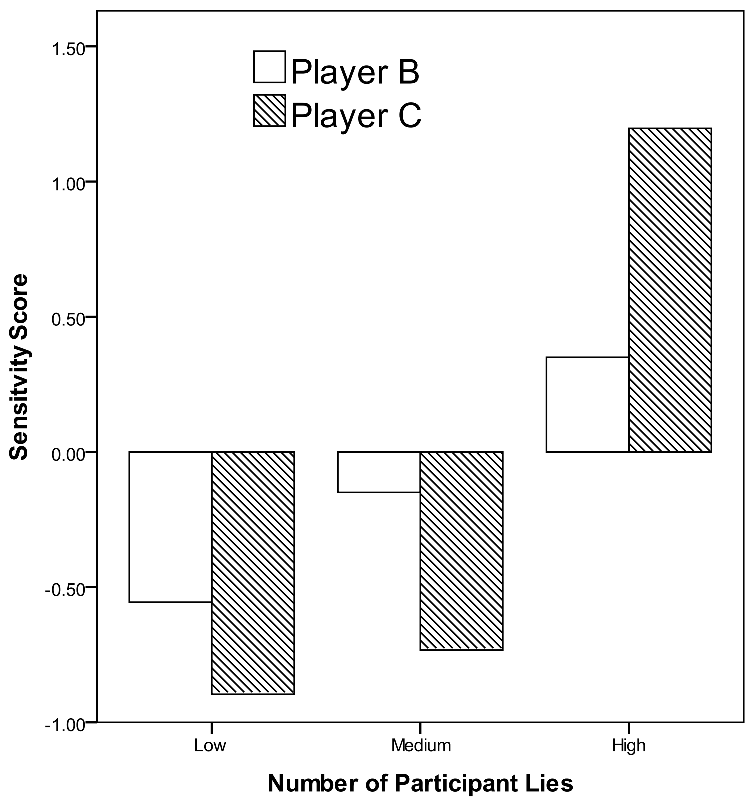

Figure 2. The interaction between Player Role and the Number of Lies by Participants on sensitivity scores.

Finally, the same (2) x 3 ANOVA was performed, but with response bias as the dependent variable. Player Role was a significant factor, $F(1, 57) = 7.22$, $p < .01$. Participants in the Player B role had higher response bias scores ($M = -.09$, $SD = .09$) than did those in the Player C role ($M = -.49$, $SD = .12$). Participants in the Player B role were thus more likely to assume that what they heard was the truth. Number of Participant Lies told was also a significant factor, $F(2, 57) = 7.08$, $p < .01$. Scheffe's post hoc tests show that groups with a high number of lies ($M = -.70$, $SD = .13$) were significantly different from groups with low ($M = -.05$, $SD = .14$) or medium numbers of lies ($M = -.13$, $SD = .13$). Participants in groups where lots of lies were told were thus less likely to assume that statements they heard were truths. The interaction between Player Role and the Number of Participant Lies on response bias scores was not significant, $F(2, 57) = 1.72$, $p > .05$.

8.0 DISCUSSION

In general, Hypothesis 1 was not supported by this study. In analyses of confederate lies, Player Role did not affect success, sensitivity, or response bias. Further analyses of success and sensitivity in the detection of player lies rather than confederate lies were conducted. In those analyses, the expected effect for Player Role was found. When participants were judging one another's' lies, those in the interdependent role had better success at detecting deception than did those in the independent role. However, this effect did not extend to sensitivity scores. Analyses of response bias on participant lies found that those in the Player C role were more likely than those in the Player B role to assume that what they heard was a lie.

Hypothesis 2 was not supported by analyses of success scores, but it was supported by analyses of sensitivity scores. As the number of lies told by the confederate increased, success decreased. However, as the number of confederate lies increased, sensitivity increased. Response bias was not affected by the number of lies that the confederate told. When participant lies were analyzed, the same pattern of results emerged; as the Number of Lies increased, success declined and sensitivity rose. These results can be explained in terms of the truth bias, which will be discussed later. Further analyses showed that response bias was significantly lower in groups with a high number of participant lies. Thus, people were less likely to assume that what they heard was the truth when a lot of lies occurred.

Tests of the third hypothesis produced mixed findings. Player Role and the Number of Lies told by the confederate did not interact to affect success scores. And no interaction effect between Player Role and the Number of Confederate Lies was found for sensitivity scores. Finally, no interaction effect was found between Player Role and the Number of Confederate Lies when it came to bias scores. But when the success scores were divided by type of success (success at detecting lies versus success at detecting truths), an interaction *was* found between the Number of Confederate Lies and Type of Success. As the Number of Confederate Lies increased, so did success at detecting lies. However, the pattern of results for success at detecting truths was more complex. When the confederate told zero lies, success at detecting truths was relatively high. But success decreased when the confederate told three lies. When the confederate told six lies, success scores regained their former level.

Regarding my research question, the results showed that the rate of lying by participants was not affected by how often the confederate lied. People may not always be encouraged to lie when another group member lies. People have different motivations to lie and may fear being caught telling lies. In my experiment, there was no real reason for the independent player to lie, so confederate lying may have had less impact on that person.

Were participants generally good or bad at detecting deception? The average success rate was 56.4%. This rate was significantly higher than 50%. Participants were thus successful more often than would be expected by chance. However, sensitivity scores were not significantly greater than zero, the chance level for such scores.

I believe that the failure to find the hypothesized effects for role may be due to my methodology. In most of Cosmides' research, participants read scenarios about hypothesized cheating, but were not cheated themselves. It is possible that the cheating detection module is activated whenever someone encounters a situation where cheating might occur - - it is not necessary for the person to be cheated personally. Thus, both players in my experiment may have had their cheating detection modules activated, because the instructions for both roles were told to both participants. The instructions informed participants that cheating was possible and thus may have activated their cheating detection modules, regardless of whether they were the player who could be cheated.

Changing the number of lies did change detection rates. As the number of lies increased, success rates went down and sensitivity rates went up. Success rates likely went down because of the truth bias. If participants did not expect to be lied to, then they would have been correct more often when the number of lies was low. Evidence for this interpretation can be found in the analyses that separated the detection rates by type of success. Participants were much more successful at detecting truths than lies. Sensitivity scores also increased as more lies were told. This supported the second hypothesis, that as more lies were told, people would become more sensitive to lies. This may mean that someone who hears many lies will be more likely to have their cheating detection module activated. Analyses of participant lies also provided support for the second hypothesis. Participants were more likely to assume that statements were lies as the Number of Participant Lies increased. However, these analyses also showed that Player B was less likely than Player C to assume that statements were lies. This may have been a result of Player B comparing the two other players. Player C may have seemed inherently more trustworthy to Player B than the confederate player, because the betting rules did not allow for Player C to cheat Player B. Thus, when comparing the statements of the two players, Player B may have been less likely to assume that Player C lied. Player C was probably less likely to compare the two players because neither player could cheat her.

The confidence levels associated with participants' judgments were moderately high. There were no significant relationships among confidence, success, and sensitivity scores. But those who were more confident had larger response bias scores - - they tended to assume that what they were told was the truth. It may be that participants assumed that everyone was telling the truth and that belief made them more confident. The weak relationship between confidence and accuracy in my experiment mirrors the results from research on eyewitness account accuracy and confidence (Sporer, Penrod, Read, & Cutler, 1995).

Several conclusions can be drawn from this research. There was preliminary evidence that people become more sensitive as more lies are told. There was little relationship between one's confidence and one's accuracy at deception detection. Finally, just being present in a situation in where cheating can occur may be enough to activate the cheating detection module; one does not need to be cheated personally.

8.1 Strengths and weaknesses

This experiment had several strengths. It was the first attempt to combine research on cheating with research on deception detection. Such a combination could provide insight into deception detection. Previous studies by Cosmides (2005) of the cheating detection module all involved scenario studies. Participants were given descriptions of situations and asked to determine when cheating could or could not occur. In my experiment, participants were actually involved in a cheating situation. I also attempted to cover new ground in deception detection. Typically, deception detection research involves viewing a recorded message and trying to determine whether the sender is lying (Bond & DePaulo, 2006). It would be useful to know how well people can detect deception in less artificial situations. In my experiment, participants did not watch a video of someone lying, but instead faced actual senders whose lies had a real impact on their outcomes.

Ironically, although the methodology of my experiment was a major strength, it was also a weakness. Finding a good way to examine deception detection and cheating together was difficult. Designing a new methodology is always challenging and potential problems are hard to foresee. The predicted difference in deception detection between Player B and Player C was not found. This was likely due to my methodology - - both roles may have activated the cheating detection module, because both players heard the rules and understood that cheating could occur.

8.2 Future directions

There are many possible future directions that my research could take. The most important would be delivering instructions and rules separately to players in different roles. If this change were implemented, then Player B's cheating detection module could be activated, without also activating the module of Player C. A methodology like this might find clearer differences in deception detection rates based on activation of the cheating detection module.

Although participants were told that they could lie, the truth bias may have led them to assume that no one was lying. Another simple change might be to add to the instructions a sentence referring to lying by past participants in similar experiments. One could even tell participants exactly how often people said that they lied in the past. This might weaken or even remove the truth bias.

My experiment looked at the impact of the number of confederate lies on confederate judgments and the impact of participant lies on participant judgments. But one could also examine how confederate lying affected judgments of participant lies, or how participant lies affected judgments of the confederate's statements.

Finally, my experiment suggests that when more lies are told, the accuracy of deception detection rises. To examine this result more carefully, experiments that involve more than ten rounds, and greater variety in the number of lies told, could be performed. For example, there may be a limit to sensitivity changes as the number of lies increases. A larger number of potential lies told might extend the results found in this experiment. Future experiments could also examine the impact of time. Does a participant who detects many lies in the first few rounds have better or worse accuracy in later rounds?

ENDNOTES

[1] Player role was treated as a repeated measures variable because both roles occurred in every group. The structure of these data was thus comparable to that of a group (or individual) that was observed twice.

[2] In later analyses, participant lies was added as a covariate and no significant changes occurred in the results reported in this paper.

BIBLIOGRAPHY

Bond, C., & DePaulo, B. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214-234.

Chen, I. (2009). Brain cells for socializing: Does an obscure nerve cell help explain what gorillas, elephants, whales – and people – have in common? *Smithsonian Magazine.* Retrieved from http://www.smithsonianmag.com/science-nature/The-Social-Brain.html

Cheng, Patricia, W., & Holyoak, Keith, J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*(4), 391-416.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276.

Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. Buss (Ed.), *Evolutionary psychology handbook* (pp. 584-627). NY: Wiley.

Cropanzano, R., & Mitchell, M. 2005. Social exchange theory: An interdisciplinary review. *Journal of Management*, *31*(6), 874-900.

Dawes, R., & Messick, D. (2000). Social dilemmas. *International Journal of Psychology, 35*(2), 111-116.

DePaulo, B., Epstein, J., & Wyer, M. (1993). Sex differences in lying: How women and men deal with the dilemma of deceit. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 126-147). New York: Guilford.

DePaulo, B., Kirkendol, S., Tang, J., & O'Brien, T. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12,* 177-202.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*(9), 913-920.

Forrest, J., & Feldman, R. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin, 26,* 118-125.

Frank, M., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stakes lies. *Journal of Personality and Social Psychology, 72*(6), 1429-1439.

Gibson, E., & Walk, R. (1960). The "visual cliff." *Scientific American, 202*, 67-71.

Hinkle, S., Taylor, L., Fox-Cardamone, L., & Crook, K., (1989). Intragroup identification and intergroup differentiation: A multicomponent approach. *British Journal of Social Psychology*, *28*, 305-317.

Homans, G. 1958. Social behavior as exchange. *American Journal of Sociology, 63*, 597-606.

Isaac, G. (1978). The food sharing behavior of protohuman hominids. *Scientific American, 238*(4), 90-108.

Jones, E., & Davis, K. From acts to dispositions: The attribution process in person perception. *In* Berkowitz, L. (Ed.), *Advances in experimental social psychology (2)*. New York: Academic Press, 1965.

Keating, P. (2005). *D-prime (signal detection) analysis.* Retrieved from http://www.linguistics.ucla.edu/faciliti/facilities/statistics/dprime.htm

Kelley, H. 1968. Interpersonal accommodation. *American Psychologist, 23*(6), 399-410.

Komorita, S. & Parks, C. (1996). *Social dilemmas.* Boulder: Westview Press.

Levine, T., Park, H., & McCornack, S. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." *Communication Monographs, 66,* 125-144.

Mealy, L., Daood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology, 17,* 119-128.

Rotter, J. (1975). Generalized expectancies for interpersonal trust. *American Psychologist, 26*(5), 443-452.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7,* 58-92.

Sporer, S., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118 (3),* 315-327.

Vanneste, S., Verplaetse, J., Van Hiel, A., & Braeckman, J. (2007). Attention bias toward noncooperative people. A dot probe classification study in cheating detection. *Evolution and Human Behavior, 28,* 272-276.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel. A kernel of truth in predictive cheating detection, *Evolution and Human Behavior*, *28,* 260-271.

Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15*, 187-203.

Weick, K. (1984). Small wins: Redefining the scale of social problems. *American Psychologist, 39*(1), 40-49.

Zuckerman, M., Koestner, R., Colella, M., & Alton, A. (1984). Anchoring in the detection of deception and leakage. *Journal of Personality and Social Psychology, 47 (2),* 301-311.