INDIVIDUAL DIFFERENCES IN DETECTING COOPERATIVE INTENT FROM EMOTIONAL FACES AT ZERO ACQUAINTANCE

by

Andrew George Siske III

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Bachelor of Science in Neuroscience with Distinction

Spring 2015

INDIVIDUAL DIFFERENCES IN DETECTING

COOPERATIVE INTENT FROM EMOTIONAL

FACES AT ZERO ACQUAINTANCE

by

Andrew George Siske III

Approved: _____

Michael Kuhlman, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Chad Forbes, Ph.D.
Committee member from the Department of Psychological and Brain Sciences

Approved: _____

Jaclyn Schwartz, Ph.D.
Committee member from the Board of Senior Thesis Readers

Approved: _____

Michael Arnold, Ph.D.
Directory, University Honors Program

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Kuhlman for his crucial guidance and support. You have been a wonderful mentor these past two years. I would also like to express thanks to my other committee members, Dr. Schwarz and Dr. Forbes, for their support and critical feedback. I also extend immense gratitude to Emily Anania, for without you I probably never would have actually written this thesis. Finally, I must give a big thank you to Adam Stivers. Your support, insight, and (especially) patience have enormously contributed to any amount of success that I may claim to have achieved.

TABLE OF CONTENTS

LIST OF FIGURES vi
ABSTRACT vii

1 INTRODUCTION 1

1.1 Social Dilemmas 1

1.1.1 Overview 1
1.1.2 The Tragedy of the Commons: Natural Resource Dilemmas 2
1.1.3 The Free Rider Problem: Public Goods Dilemmas 3

1.2 Solving Social Dilemmas 4

1.2.1 Structural Solutions 4
1.2.2 Cooperation in the Absence of Authority 5

1.3 The "Morality" Solution 6

1.3.1 People Differ in Social Preferences, or Goals for Cooperation 6

1.3.1.1 Measuring Social Value Orientation (SVO) 7

1.3.2 Cooperation Can Be Induced or Discouraged 8
1.3.3 People Need Information to Form Expectations 10
1.3.4 Nonverbal Cues Can Signal Cooperativeness 12

1.4 Theory of Mind 14

2 METHODS & MEASUREMENTS 16

2.1 Participants 16
2.2 Procedure Overview 16
2.3 Social Rating Task 17
2.4 Reading the Mind in Eyes 19

3 RESULTS 21

3.1 Internal Consistency of Reading the Mind in the Eyes (RME) Measure 21
3.2 Measuring SVO Detection Accuracy 21
3.3 Main Effects for Accuracy 22

3.3.1 Relation Between Theory of Mind Score and Between Subject Variables 22

3.3.2 Analyses of Accuracy 22

3.3.2.1 Accuracy and the Between-Subject Variables 23
3.3.2.2 Accuracy and Theory of Mind 23
3.3.2.3 Accuracy as a Function of Between-Subject and Within-Subject Effects 24

3.3.2.3.1 Effects for Sex of Photo 25
3.3.2.3.2 Effects for SVO of Photo 26
3.3.2.3.3 The Sex of Photo by SVO of Photo Interaction 26

3.3.2.4 Accuracy for Each Sex by Theory of Mind 28

4 DISCUSSION 29

4.1 General Overview 29
4.2 Theory of Mind 29
4.3 Results Not Related to Theory of Mind 34

REFERENCES 37
APPENDIX

SOCIAL RATING TASK INSTRUCTIONS 43

LIST OF FIGURES

Figure 1. An example of a "social decision task" like that shown in the Social Rating Task instructions 18

Figure 2. Example of a stimulus from the Reading the Mind in the Eyes task. Taken from the Qualtrics survey actually used in the experiment 20

Figure 3. Accuracy as predicted from Theory of Mind scores 24

Figure 4. Means for the Sex of Photo by SVO of Participant interaction 25

Figure 5. Means for the Sex of Photo by SVO of Photo interaction 27

Figure 6. Cooperation Likelihood for the Sex of Photo by SVO of Photo Interaction 28

ABSTRACT

In social situations of conflict, individuals differ in their tendency to behave cooperatively, individualistically, and competitively. The limits on consumption imposed by the general scarcity of resources often leads people to avoid the short run costs imposed by cooperation, opting instead to maximize their own benefit at some cost to the collective's welfare. If every individual behaves in this selfish manner, however, all are left worse off in the long run than if all had acted cooperatively. It is thus of interest to investigate how cooperative people may identify other cooperative people in order to achieve superior outcomes, and to avoid exploitation. Previous research suggests that cooperative / non-cooperative predispositions are detectable from minimal information, including conversational mannerisms and emotional expressions. We investigate the role Theory of Mind (ToM), the ability to perceive and understand the mental states of others, might play in the ability to recognize such predispositions. We hypothesized that those better at identifying emotional states from a pair of eyes (i.e. Baron-Cohen's "Reading the Mind in the Eyes" task) would more successfully (accurately) recognize the cooperativeness of an unknown individual. Our results indicate that a relationship exists between ToM and cooperativeness detection. Contrary to our expectations, however, this relationship was found to be negative. Given that our notion of cooperativeness comes from research on social dilemmas, the current work offers an overview of social dilemmas and how social dilemmas are "solved." We conclude with a discussion surrounding the implications of our results for theoretical and real-world social dilemmas.

Chapter 1

INTRODUCTION

1.1 Social Dilemmas

1.1.1 Overview

A social dilemma is a situation in which at least two individuals must choose between (a) non-cooperation, or maximizing his/her own benefit at some cost to collective benefit, and (b) cooperation, or maximizing collective benefit at some cost to his/her own (Dawes, 1980; Komorita & Parks, 1995; Shelley et al., 2010). Others have characterized social dilemmas as involving a conflict between individual rationality and collective rationality (e.g. Kollock, 1998).

Based on the rational choice theory assumption of universal human egocentrism (commonly known as *homo economicus*), non-cooperation should be the most likely in social dilemmas, given that non-cooperation dominates cooperation. Importantly, this suggests that social dilemmas will consistently result in mutual non-cooperation. However, the result of universal non-cooperation is known in game theory, is a *deficient equilibrium* (Kollock, 1998). This is to say that there exists at least one other outcome that is a *Pareto improvement* – an alternative that would yield higher payoffs to at least one person while leaving none worse off. At the extreme of universal mutual cooperation, a *Pareto efficient* outcome occurs. This is situation in

which no one individual can be made better off without making another worse off. This conflict between individual rationality and collective outcome constitutes the essence of a social dilemma. The universal pursuit of (individually rational) self-interest produces an outcome which is worse for all parties than the outcome produced by the universal pursuit of (collectively rational) cooperation

1.1.2 The Tragedy of the Commons: Natural Resource Dilemmas

To give an example of a social dilemma: Imagine a small village wherein the inhabitants get much of their nutrients from eating fish that are caught in the local lake. Each member of the community faces a daily decision: how many fish to catch today. Each individual clearly sees that having more fish is better than having less fish. More fish equals more food, and when weather is bad, having a few extra fish stored away would be convenient. Thus, each has incentive to catch as many fish as they can manage. If each member of the village acted in this way, however, the rate at which the fish are harvested would likely exceed their reproduction rate, and the village lake's supply of fish would eventually deplete. Hence, when each individual in this village acts in his/her own self-interest (i.e. capture as many fish as possible), all individuals will be left in the worst possible condition (i.e. there will be no more fish to catch and eat). This is an example of a common-pool resource dilemma (Hardin, 1968; Parks, Joireman, and Van Lange, 2013). In such situations, the dilemma lies in how a common resource (such forests, fish, or energy) will be over-consumed when individual rationality ("common sense") compels behavior.

1.1.3 The Free Rider Problem: Public Goods Dilemmas

Another type of social dilemma is a public goods dilemma. In these situations, a good is provided for unrestricted consumption by all who so choose to consume it, but its existence depends a certain level of contribution (in money, time, effort) from its users. It is easy to imagine that most users will not necessarily feel compelled to contribute despite their consumption of it, as they believe others will contribute enough such that they can enjoy the good without paying for it. Consider how many people who listen to National Public Radio but do not contribute funds to it, or how many citizens who exist within democratic societies do not vote during political elections (Van Lange et al., 2013). In the context of a public good, non-cooperation is choosing not to contribute to provision of a good one consumes freely. Cooperation in such a situation would be to contribute some amount to the maintenance of the good. However, by not contributing to a public good, an individual maximizes his/her own benefit, without sacrificing anything in return. Thus public goods dilemmas create an incentive to "free ride". But, the consequence of universal free riding is that the public good is not provided. In this case everyone is worse off than if they had all cooperated. Yet another example of a social dilemma.

A vast body of literature dedicated to defining, understanding, and solving social dilemmas exists (Dawes, 1980; Van Lange et al., 2013). Of most interest to the current project is how social dilemmas are solved in the absence of strong external structural forces that coerce cooperation, such as taxes, and laws that sanction self-interested behavior. Citing Malthusian population growth theory, Hardin specifically addresses the tragedy of overpopulation ("freedom to breed"), a social dilemma that he purports to have "no technical solution" (Hardin, 1968). By this, he meant that he

could imagine no viable structural solution contained within the natural sciences or in the progression of technology. Rather, a solution to this dilemma would have to emerge from an extension of morality; a change in psychology. Here Hardin is making an important distinction between ways in which dilemmas might be solved; by structural versus psychological processes.

1.2 Solving Social Dilemmas

1.2.1 Structural Solutions

Broadly speaking, there are two kinds of potential solutions to a social dilemma. The first of the two classes of solutions can appropriately be called *structural solutions*. These are solutions resulting from changes made to the system surrounding the social dilemma in question. Structural solutions require that an entity exists that is powerful enough to make such changes – usually a federal or state-level governmental agency, or perhaps a civic entity within which members can vote for structural modifications. Structural changes to a dilemma scenario can consist of enforcing sanctions for non-cooperators, offering rewards for cooperators, implementing means of communication between users within the structure, and, in the absence of an established central government, appointing an individual involved in the dilemma to a position of leadership and power (Parks, Joireman, and Van Lange, 2013).

The structural changes listed above have been demonstrated to successfully promote cooperation in experimental social dilemmas (Guererk, Irlenbusch, & Rockenbach, 2006; Balliet, Mulder & Van Lange, 2011; Bowles and Gintis, 2011). It is reassuring that this is so; there is evidence we are to some extent able to improve

"tragic" situations of natural resource allocation and public good funding. As already mentioned, however, such solutions usually require the intervention and oversight of a third party (i.e. government officials). In order for a structural solution to be successful, the party implementing changes and enforcing sanctions must be regarded as legitimate (Ertan et al., 2009). Considering the ideological tensions surrounding the role of government in the U.S. especially and in Western countries generally, it is clear that structural solutions to dilemmas importantly and controversially impact the daily lives of individuals. Coercion applied by a third party may not always be successful in motivating individuals to cooperate rather than defect. Some people have strong ideological preferences for minimal third party intervention in their daily affairs. Additionally, the economic costs and bureaucratic institutions surrounding structural solutions may reduce their efficiency and thus attractiveness.

Is there another way? Could cooperation be promoted and maintained in the absence of externally imposed sanctions?

1.2.2 Cooperation in the Absence of Authority

Pruitt & Kimmel (1977) produced a theory of cooperation years ago called Goal / Expectation Theory (GET henceforth). They hold that an individual will cooperate if both of two criteria are met: (1) s/he has the *goal* of cooperating, and (2) s/he *expects* that those around him/her will cooperate. As seen above, structural modifications work. This success can be understood through the lens of GET. Structural changes make cooperation *the* objective goal for all, even if it is not necessarily each person's spontaneous subjective goal. More importantly, a

sanctioning system makes it more reasonable to *expect* that others will cooperate – if they fail to do so, they will be thrown in jail, fined, or otherwise sanctioned.

However, it may not be necessary to make structural changes to a social dilemma in order to induce cooperation. According to GET, we need only have individuals with both the goal of cooperating and an expectation that this goal is shared for cooperation to emerge. It may be that individuals with the goal of cooperating exist, and it may also be the case that there are ways to communicate this cooperative intent (and hence create the expectation of cooperation). Let us start with the former. Is there evidence that some people have the goal of cooperating with others?

1.3 The "Morality" Solution

1.3.1 People Differ in Social Preferences, or Goals for Cooperation

There is empirical evidence that people are not as selfish as rational choice theory would have us believe (Kuhlman & Marshello, 1975; Murphy, Ackermann, and Handgraaf, 2011). In laboratory studies of social dilemmas, in which participants freely choose between cooperation and noncooperation in the absence of external sanctions, it has been found that a surprisingly high number of people behave in a cooperative manner (Dawes, 1980; Murphy & Ackermann, 2014).

Again, GET theory is relevant. A large amount of literature has investigated and validated the existence of differing social preferences. One of the most heavily researched and best validated personality measures is one's *Social Value Orientation* (SVO), which categorizes individuals based on the relative value the individual assigns to individual payoffs and to the payoff of an interdependent other (Kuhlman &

Marshello, 1975; Balliet, Parks, and Joireman, 2009). Much of the social dilemma literature focuses on three distinct SVOs: Cooperative (J), or concern for *joint* gains; Individualist (O), or concern for *own* gain while indifferent to another's gains; and Competitive (R), or concern for *relative* gain (i.e. maximizing own gains over other's gains). Importantly, some 50% of people measured act as Cooperators (Au & Kwong, 2004).

1.3.1.1 Measuring Social Value Orientation (SVO)

Messick & McClintock (1968) were the first to use decomposed games to measure differences in social motivation between people. These researchers identified and demonstrated the existence of Cooperators, Individualists and Competitors using a very simple decision task known as a decomposed game. An individual playing one of their decomposed games is assigned the task of choosing between two allocations of imaginary money to themselves and another person. As one example, Allocation A consists of giving one's self $50 and the other $50; Allocation B consists of giving one's self $60 and the other $35. Selecting Allocation A suggests that this individual has a Cooperative SVO, whereas selecting Allocation B suggests s/he has a either an Individualistic or Competitive SVO. A distinction between Individualistic and Competitive SVO's can be made by having the participant play another game in which Allocation A is $60 for self and $40 for the other and Allocation B is $50 for self and $0 for the other. Individualists would choose A and Competitors would choose B. In such a task, there is no strategy involved; players must simply choose which of the available allocation packages they deem most desirable. The participant's SVO is thus said to be "revealed by his/her preferences."

In the current project, participants' SVO was assessed with a widely used decomposed game technique known as the Ring Measure (Liebrand, 1984). This, and other decomposed game techniques for measuring SVO have been used around the world since the 1970's (see Murphy, Ackermann, and Handgraaf, 2011 for a good SVO measurement overview).

An important finding is that regardless of country or language, the SVO of Cooperation is almost always the modal group (50% or higher), followed by Individualism (around 30%) and Competition (around 12%) (Murphy, 2014). It appears that there is reason to assume that more people in general have cooperative as opposed to non-cooperative motivations, which corresponds to the Goal aspect of GET theory. Using different procedures, research in behavioral economics (Bowles & Gintis, 2011) has come to the conclusion that we can be regarded as a cooperative species. That most people in the world have cooperative goals is comforting, but it can be shown that having such a goal is not sufficient to maintain cooperation. This corresponds to the Expectation component of GET.

1.3.2 Cooperation Can Be Induced or Discouraged

It seems that a large number of people regard gains for others as "good." Although an individual's SVO has been shown to be consistent over time (Bogaert et al., 2008; Murphy, 2014), it has been shown that people of each orientation sometimes modify their "default" behavior in response to decisions made by others within a relationship. A 1975 study by Kuhlman and Marshello had participants of various SVOs engage in sequential (repeated) experimental two-person social dilemmas with computers preprogrammed to behave a certain way. Specifically, some participants

were paired with preprogrammed *unconditional cooperators*, who would choose the cooperative option in every interaction regardless of the participant's previous decisions; some were paired with *unconditional defectors*, who would choose the defective option in every interaction regardless of the participant's decision history; still others were paired with computers using a *tit-for-tat* strategy, in which the first decision is always cooperative, and all following decisions reciprocate the decision made by the participant.

It was shown that Cooperative SVO's always cooperated with unconditionally cooperative and with tit-for-tat utilizing players, but always defected against unconditional defectors. Competitors always defected, regardless of partner (as SVO theory predicts). Individualists always defected against unconditional cooperators and unconditional defectors, but cooperated with partners using tit-for-tat. In another work, it was shown that even Competitors can be induced to cooperate (Sheldon, 1999). Thus, given the proper conditions, Cooperators may be prone to defection, and competitors may be prone to cooperation.

In Kuhlman & Marshello (1975), Cooperators became non-cooperative when they came to *expect* that cooperation would not be reciprocated (i.e. when playing with unconditional defectors). Similarly, Individualists became cooperative when it was expected that non-cooperation would lead to lower outcomes (i.e. when playing with tit-for-tat). It should be obvious that *expectations matter a lot* for individuals in social dilemmas. In the current project, we are primarily concerned with how one comes to expect that those around him/her will act cooperatively.

1.3.3 People Need Information to Form Expectations

In the previously discussed Kuhlman & Marshello (1975) paper, the participants' expectations about their respective other players emerged out of repeated play – the history of interactions served as information to how future interactions may turn out. In other words, history served a communicative function of intent. In many real world dilemmas, however, repeated interactions with partners do not necessarily occur, depending on the context of the dilemma.

Previous research has shown that communication leads to substantial increases in cooperation (Balliet, 2010). Much of this research has focused exclusively on verbal communication. The typically studied mediums of verbal communication are face-to-face (Orbell, van de Kragt, and Dawes, 1988; Kerr & Kaufmann-Gilliland, 1994) and written messages (Duffy & Feltovich, 2002; Tazelaar, Van Lange, and Ouwerkerk, 2004).

Nonverbal communication has been studied in the context of social dilemmas as well, but less extensively. Haley and Fessler (2005) demonstrate that the presence of a set of stylized eyes on the desktop background of the participant's computer increased generosity (cooperativeness) compared to controls in a decomposed game. Similarly, other studies have found that participants exposed to a set of eyes contribute more money to community drink funds (Bateson, Nettle, and Roberts, 2006) and are more likely to go out and vote (Panagopoulos, 2014). Another study showed that participants were more generous in a laboratory public goods dilemma following a mutual eye gaze with another participant, a gentle touch from another participant, and written communication with other participants, but not after keeping a rhythm with other participants (Kurzban, 2001).

It seems that the mere presence of a pair of (non-real) eyes, not to mention a seemingly informative mutual eye gaze, suffices to increase cooperation. Such results may be interpreted in very interesting ways, but the real world applicability of this research is nonetheless limited. It will not typically be the case that a picture of eyes will be around to induce cooperation, nor will it always be the case that locking eyes with an interdependent other will signal cooperative intentions. In Kurzban (2001), the participants who exchanged the mutual eye gaze inherently had more information about their dilemma than most real world dilemmas offer. In such settings, expectations often have to be formed from minimal information.

Imagine walking down a city street a few blocks away from your apartment. A male stranger taps your shoulder from behind: "Pardon me, but may I use your cell phone for just a moment? Mine is dead and I have no way of reaching my sister who is surely concerned about me. This is my first time to the city, and I do not know what else I can do." You are placed in a dilemma: Do you lend this stranger your phone, trusting that he will not simply take off running when your guard appears down? Or do you reject his possibly genuine plea for help, briskly walk away and continue about your business as planned?

In such a scenario, it is impossible to know how this individual has conducted himself before. Perhaps he is a thief employing a consistently successful tactic. Perhaps he is true to his word and is lost, and has every intention of returning your phone immediately after one call. It is also difficult to trust his word. Sure, he may verbally indicate that he will not steal your phone, but his words may not reflect his intended behavior.

In such a situation, it would be critically helpful if something about the stranger's social value system were discernable from "something about him" – from his mannerisms, or even from his expressions of emotion. Is there any evidence that such minimal information could suffice for making an accurate judgment about an individual's cooperativeness?

1.3.4 Nonverbal Cues Can Signal Cooperativeness

Previous research has shown that prosocial behavioral predispositions are communicated through nonverbal cues. Schug et al. (2010) filmed participants while engaged in a laboratory social dilemma game, and compared the facial expressions of individuals who acted cooperatively versus those who acted non-cooperatively. It was found that cooperative participants were significantly more emotionally expressive than non-cooperators, i.e. displayed emotional expressions more frequently, regardless of valence.

Other work using video-clips of self-reported altruists and self-reported non-altruists as stimuli has found that altruists were rated as more helpful (Brown, Palameta, and Moore, 2003; Oda et al., 2009), more altruistic (Oda et al., 2009), and were entrusted with more money in an experimental game (Oda, Naganawa, Yamauchi, et al., 2009). Additionally, Verplaetse, Vanneste, and Braeckman (2007) took pictures of individuals (1) before playing an experimental dilemma game, (2) during a decision-making moment in a practice round, and (3) during a decision-making moment in a real round. It was found that participants were able to discern cooperators from non-cooperators above chance, but only from the picture taken during the real round.

Shelley et al. (2010) filmed Cooperators, Individualists, and Competitors discussing the events of the previous day. Using this footage as stimuli, it was found that ProSocials (Cooperative SVO) displayed more "enjoyment smiles" than ProSelfs (Individualists and Competitors combined) (Study 1) and participants (judges) were able to distinguish ProSocials from ProSelfs at a level above chance (Study 2). It was also found that females were more likely to be judged as cooperative, whereas males were more likely to be judged as non-cooperative.

In another study reported in the same paper, participants viewed still photos of individuals (who were different from those showed to participants in Study 2) posing neutral expressions (Study 3), happy expressions (Study 4), and a gallery (happy, angry, sad, afraid, disgusted, surprised, and friendly) of emotional expressions (Study 5). Participants were able to distinguish ProSocials from ProSelfs in studies 4 and 5 (using emotional faces) but not in Study 3 (using neutral faces). The results of Study 4 were replicated in Page (2012). For both Shelley et al. (2010) and Page (2012), the SVO of the judge did not significantly influence SVO detection accuracy.

From the work mentioned above (especially Shelley et al., 2010), two important conclusions can be drawn: (1) Information about an individual's Social Value Orientation is available from his/her conversational mannerisms and even from an emotional photographs, and (2) People are sensitive to such information, regardless of their Social Value Orientation.

The current project further investigates conclusion (2) and extends Shelley and Page's research by asking the following: Are there individual differences in the ability to detect the cooperativeness of an unknown other from an emotional photograph?

And if so, what might at least partially explain these differences? It is proposed that an individual's Theory of Mind plays a role in SVO detection accuracy.

1.4 Theory of Mind

Theory of Mind (ToM) refers to the ability to "impute mental states" to others (Premack &Woodruff, 1978); to properly understand and interpret what others are thinking, feeling, or desiring (Sylwester et al., 2012). Researchers often distinguish cognitive ToM from affective ToM, which have been called *perspective-taking* and *empathy*, respectively (Paal & Bereczkei, 2007; Bodden et al., 2010). That being said, ToM has been conceptualized and measured in many ways, resulting in a rather incoherent literature (Apperly, 2012; Bosco et al., 2014).

Much research on ToM has focused on its development in children (Walker & Murachver, 2012; Lillard et al., 2013; Slaughter et al., 2015) and on deficiencies in clinical samples (Baron-Cohen et al, 2001; Richell et al, 2003; Tella et al., 2015; Xi et al., 2015). Children do not develop a Theory of Mind until around age 5 (Singer & Fehr, 2005), and patients with developmental and various psychiatric disorders have shown deficiencies in ToM (particularly those with Autism Spectrum Disorders) (Baron-Cohen at al, 2001). Recent research has also studied "normal" adult populations including age differences (Henry et al., 2013; Wang & Su, 2013), correlates with personality traits (Paal & Bereczkei, 2007; Hünefeldt, Laghi, and Ortu, 2013), and correlates with various types of intelligence (Peterson & Miller, 2012; Baker et al., 2014). Of particular interest to the current research, others have studied ToM in the context of pro-sociality (Declerck & Bogaert, 2008; Sylwester et al., 2012; Van Doesum, Van Lange, and Van Lange, 2013). It has been found that individuals

with a Cooperative SVO are generally better at ToM tasks, though the effect size is quite small. (Declerck & Bogaert, 2008; Van Doesum, Van Lange, and Van Lange, 2013).

The degree to which an individual can accurately attribute beliefs, desires, and intentions to others will largely impact how s/he behaves in a given situation. Intuitively, how one perceives and understands the mental states of others importantly influences all social interactions, especially those involving a conflict of interests (i.e. social dilemmas). In other words, the extent to which an individual can properly infer the mental states of others will influence the *expectations* s/he will have about how others will behave. It is this logic that compels an investigation of the role Theory of Mind might play in how well an individual can accurately discern the cooperativeness of another person. The current research utilizes a ToM task wherein participants must discern an individual's emotional state from a photo cropped around the eyes (the task is elaborated upon in the Methods section). This measure seems to be appropriate given that we are asking participants to infer the SVO of an individual from an emotional picture (albeit one that shows the full face).

The current project hypothesizes that those who perform better on a Theory of Mind task will be more successful in detecting the cooperativeness of strangers. Specifically, it is proposed that those more successful in identifying the emotion conveyed by a pair of eyes will be more successful in detecting the cooperativeness of an individual from an emotional photograph.

Chapter 2

METHODS & MEASUREMENTS

2.1 Participants

One hundred sixty eight students enrolled in Introduction to Psychology satisfying a research participation requirement participated in the study. Of these 13 students were excluded for final analysis (reasons given in the Results Section), leaving a total of 155 (97 female, 58 male). Experimental sessions were run in small groups of 8 to 12 students at a time. Students' SVOs were measured in pretesting using Liebrand's Ring Measure (Liebrand, 1984). We selected participants such that we would have approximately equal numbers of male J's, O's, and R's, and female J's, O's, and R's. In this study, and in keeping with highly common practice in SVO research, participants were classified as either ProSocials (Cooperative SVO) or ProSelfs (Individualistic or Competitive SVO's). This is based on the fact that most of the time Individualists and Competitors do not differ; this is especially true for the non-verbal research on SVO communication and detection described above.

2.2 Procedure Overview

The experiment was conducted over the course of four days in our computer lab. Each participant completed the experiment at his or her own computing station. All tasks in each session were administered via Qualtrics. The experiment consisted of three tasks: – (1) The Social Rating Task (2) Theory of Mind assessment, and (3) Social Mindfulness assessment. Results for Social Mindfulness will not be reported in

this thesis. Without exception, each participant completed the Social Rating Task first followed by the ToM task.

Participants in each of the conditions for Social Rating Task viewed 61 photographs of still, emotional faces taken of participants during a previous study. From this previous work, we know the Social Value Orientation of each of the shown faces. Those who were viewed gave consent for their photographs to be used in research of the type conducted here.

2.3 Social Rating Task

The Social Rating Task had participants view 61 photographs of still faces expressing a particular emotion and, following previous work (Shelley et al, 2010; Page, 2012), asked participants to rate the likelihood that each viewed face would follow or utilize each of three "choice rules." The instructions for the Social Rating Task (Appendix) described these choice rules.

In the instructions, participants were shown a "social decision task" in which a choice must be made between one of three different allocations of points to yourself and an unknown "other" (Figure 1). Participants were told that in this task, the pictured player had to decide between allocation bundles A, B, and C, where A gets the player and the other each 70 points; B gives the player 100 points and the other none; and C gives the player 70 points and takes away 70 points from the other.

	A	B	C
You Get	70	100	70
Other Gets	70	0	-70

Figure 1. An example of a "social decision task" like that shown in the Social Rating Task instructions. Self(+), Other(+) would select allocation bundle A; Self(+),Other(0) would select bundle B; Self(+),Other(-) would select bundle C.

Participants were told that previous research showed that people make their choices in such situations based upon different choice rules. We described three different choice rules in terms of an individual's concern for his/her own outcome and for that of the other decision-maker: Self(+),Other(+), Self(+),Other(0), and Self(+),Other(-).

These three "choice rules" are given as stand-ins for the three primary Social Value Orientations. Self(+),Other(+) is meant to reflect the concern of a Cooperator, who values the outcome both for his/herself and for the other. Self(+),Other(0) reflects an Individualist's value orientation; concern for his/herself but indifference for the outcome of the other. Self(+),Other(-) corresponds to a Competitive orientation, in which the outcome yielding the greatest relative difference in points is desirable.

As in the previous work of Shelley et al (2010) and Page (2012) we used these choice rules in order to avoid using terms such as "cooperative" or "competitive" which could arguably create a social desirability bias. Before proceeding to the actual rating task, participants were given two quizzes to assure that they understood the choice rules. If a quiz question was answered incorrectly, the

participant was directed back to the instructions page to once again review the choice rules. The first quiz was given before an initial viewing of all 61 photographs to check for acquaintance with any of the viewed persons; the second quiz was given immediately after this viewing. Participants who answered more than two quiz questions incorrectly were excluded from data analysis.

Participants were divided into one of three experimental conditions – 'Angry', 'Happy', or 'Both' – which determined the set of photographs viewed. In the Angry condition, participants viewed 61 photographs of individuals asked to display an angry face. Similarly, those in the Happy condition viewed 61 photographs of the same individuals asked to show a happy face. Those in the Both condition viewed the angry and happy photo for each of the same 61 faces. In this condition the photographs were displayed side-by-side.

After completing the quizzes and viewing all 61 relevant photographs to screen for familiarity with the faces, participants completed the actual Social Rating Task. Participants viewed each photo (though two are viewed in the Both condition, statements relevant to all three conditions will be discussed in the singular) individually. For each displayed face, they were asked to rate the likelihood that the viewed person used each of the three choice rules. Likelihood was rated on a 7-point Likert scale in which 1 corresponded to "Very Unlikely" and 7 to "Very Likely." No labels were assigned to ratings of 2 through 6.

2.4 Reading the Mind in Eyes

To assess participants' Theory of Mind, Baron-Cohen's Revised Reading the Mind in the Eyes (RME) test (Baron-Cohen et al, 2001) was used. In this task,

participants view a series of 37 faces (cropped so as to only display their gaze) expressing a complex emotion and are asked to select (from four multiple choice answers) which emotional state best describes what the shown person is experiencing (Figure 2). The first set of eyes viewed was the same for all participants, which served as a "practice" trial; the remaining 36 eyes were randomized. Only the 36 randomized stimuli were scored. A correct answer was scored as 1; an incorrect answer as 0. The task is scored by summing up how many emotional states were correctly identified, such that a higher score corresponds to having a better Theory of Mind ability.

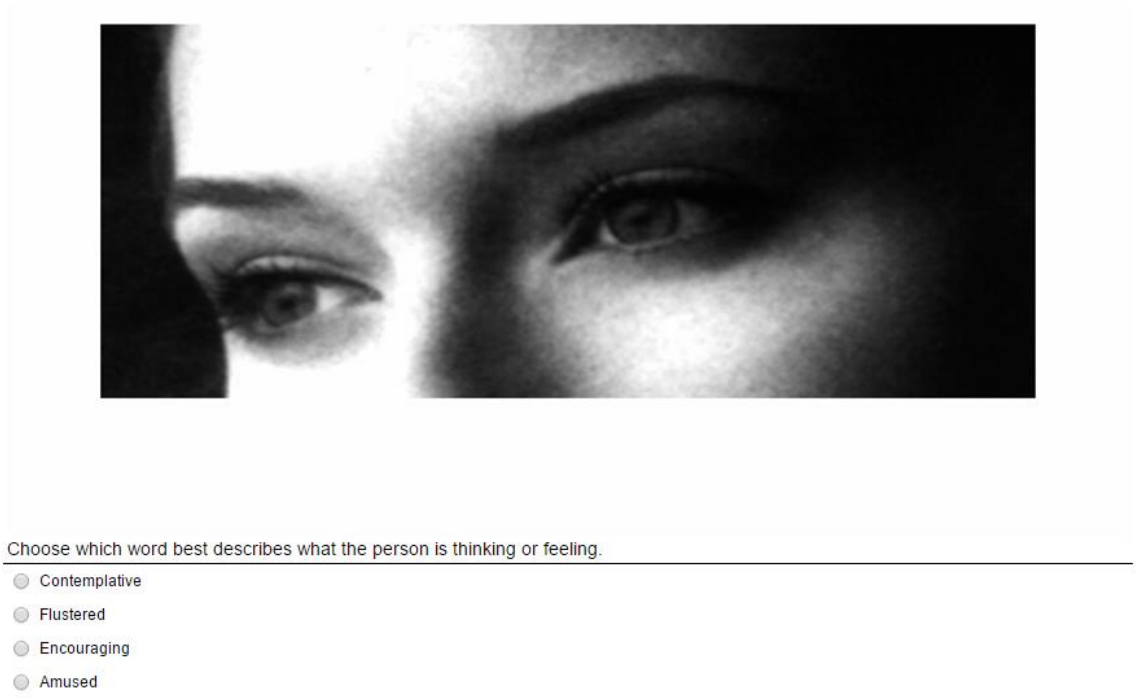

Figure 2. Example of a stimulus from the Reading the Mind in the Eyes task. Taken from the Qualtrics survey actually used in the experiment.

Chapter 3

RESULTS

3.1 Internal Consistency of Reading the Mind in the Eyes (RME) Measure

For all 155 participants Cronbach's $\alpha = 0.584$, which was lower than the generally accepted value (0.7) for a reliable measure. For male participants $\alpha = 0.729$, but for females it was only 0.36. Thus, results reported below for RME should be interpreted with caution.

3.2 Measuring SVO Detection Accuracy

Our measure of Social Value Orientation detection accuracy ("Accuracy" henceforth), was calculated as follows. Let A be the likelihood rating for Self(+), Other(+); B is the rating for Self(+), Other(0) variable B; and C is the rating for Self(+), Other(-) variable C. Thus, A, B, and C correspond to likelihood ratings for Cooperation, Individualism and Competition, respectively.

If the rated photo was that of a Cooperator we subtracted the average of the Individualism/Competition ratings from the Cooperation rating.

$$\text{Accuracy}_{\text{social}} = A - (B+C)/2$$

Thus, if the viewed individual was a Cooperator, accuracy corresponds to the difference between ProSocial and ProSelf ratings. Positive values indicate accuracy, and negative values indicate "anti-accuracy".

If the viewed individual was that of an Individualist or a Competitor, subtracted the participant's rating for being ProSocial from that of being ProSelf, according to this formula:

$$\text{Accuracy}_{\text{self}} = (B+C)/2 - A$$

Again, positive values indicate accuracy, and negative values indicate “anti-accuracy”.

3.3 Main Effects for Accuracy

To be included for analysis, the participant had to satisfy two criteria: (1) a choice consistency index on the Ring Measure of at least 0.6 and (2) no more than 2 mistakes on the quiz to measure understanding of the rating task. There were 155 participants who met both criteria.

3.3.1 Relation Between Theory of Mind Score and Between Subject Variables

In the first analysis the Theory of Mind score was the dependent variable in a 2 (Sex of Participant) by 2 (SVO of Participant: ProSocial or ProSelf) by 3 (Type of Photo: Angry, Happy, Both) factorial design. The test for the Overall Model was not significant ($F(11,141) = 0.81$, $p = 0.621$, $R^2 = 0.06$). Tests on the 11 single-df contrasts for the main effects and interaction effects were all non-significant, providing further evidence for the lack of relationship between Theory of Mind score and participant Sex, SVO or type of photo.

3.3.2 Analyses of Accuracy

Each participant judged the SVO (ProSocial or ProSelf) of 61 photos. For each photo, accuracy was computed according to the procedure described above (section 3.2). The average of these 61 accuracy scores served as the dependent (criterion) variable in the series of analyses reported below.

3.3.2.1 Accuracy and the Between-Subject Variables

Accuracy was the dependent variable in a Participant Sex by Participant SVO by Photo Type analysis of variance. The only effect that approached significance was for the Constant (Mn = 0.0354) ($F(1,141) = 3.07$, $p = 0.08$). That the constant was greater than zero suggests a slight amount of accuracy, which is consistent with the previous work of SVO judgment based on photos. The absence of differences in accuracy for all Between Subject effects is also consistent with previous research.

3.3.2.2 Accuracy and Theory of Mind

The initial analysis was a bivariate regression in which Accuracy was predicted from Theory of Mind (centered around its mean). The regression constant ($\beta_0 = 0.04$) was significantly different from zero ($F(1,151) = 4.656$, $p = 0.032$) indicating a small level of accuracy. The Theory of Mind score was significantly related to Accuracy ($F(1,151) = 5.01$, $p = 0.027$), but as can be seen in Figure 3 the relationship was negative.

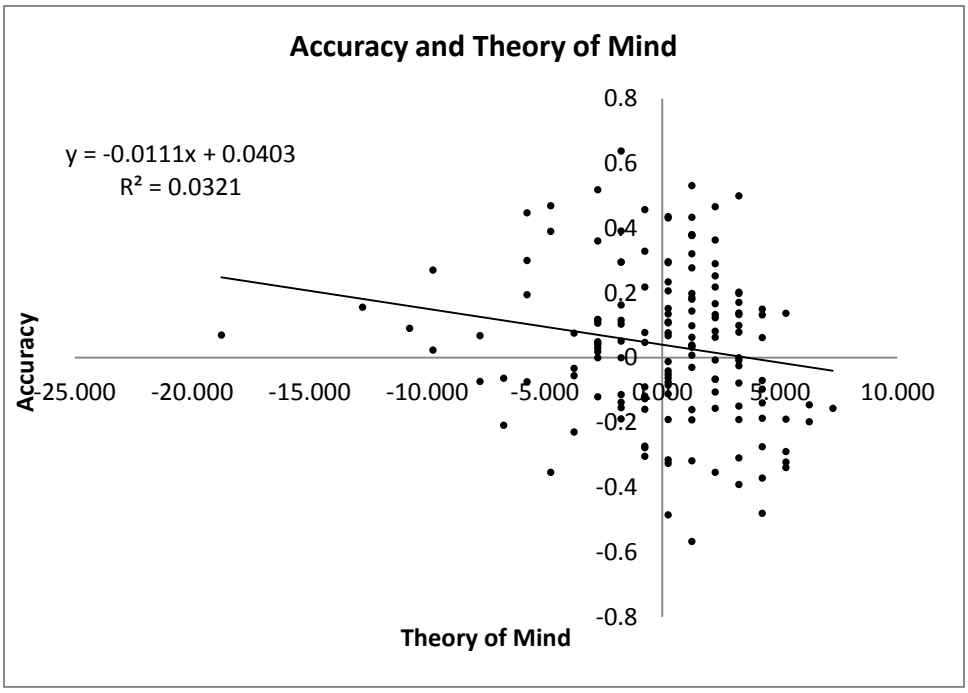

Figure 3. Accuracy as predicted from Theory of Mind scores.

Next, a series of Multiple Regressions was run to see if the Theory of Mind/Accuracy relation was moderated by the Participant's Sex, SVO, and by the Type of Photo. In each analysis no moderating effects were found. For Participant Sex by Theory of Mind $p = 0.411$, for Participant SVO by Theory of Mind $p = 0.284$, for Both Emotion Photos versus One Emotion Photos, $p = 0.65$, and for Happy versus Angry Photos $p = 0.61$. Thus, the negative relationship between Theory of Mind and Accuracy appears to be quite robust.

3.3.2.3 Accuracy as a Function of Between-Subject and Within-Subject Effects

To further explore how accuracy might have related to variables other than Theory of Mind, we performed a factorial Analysis of Variance in which Sex of Participant (2), SVO of Participant (2) Type of Photo (3), Sex of Photo(2) and SVO of

Photo (2) were the independent variables. Sex of Photo and SVO of Photo were repeated measures variables, and the other three were between subjects.

3.3.2.3.1 Effects for Sex of Photo

There was a marginal effect ($F(1,141) = 3.72$, $p = 0.056$) for the Sex of the Photo. Female Photos (Mn = 0.119), were rated more accurately than Male Photos (Mn = -0.041). This effect was moderated by the SVO of the Participant ($F(1,141) = 9.15$, $p = 0.003$). As can be seen in Figure 4, the difference between Male and Female photos was more pronounced in ProSocial than in ProSelf Participants.

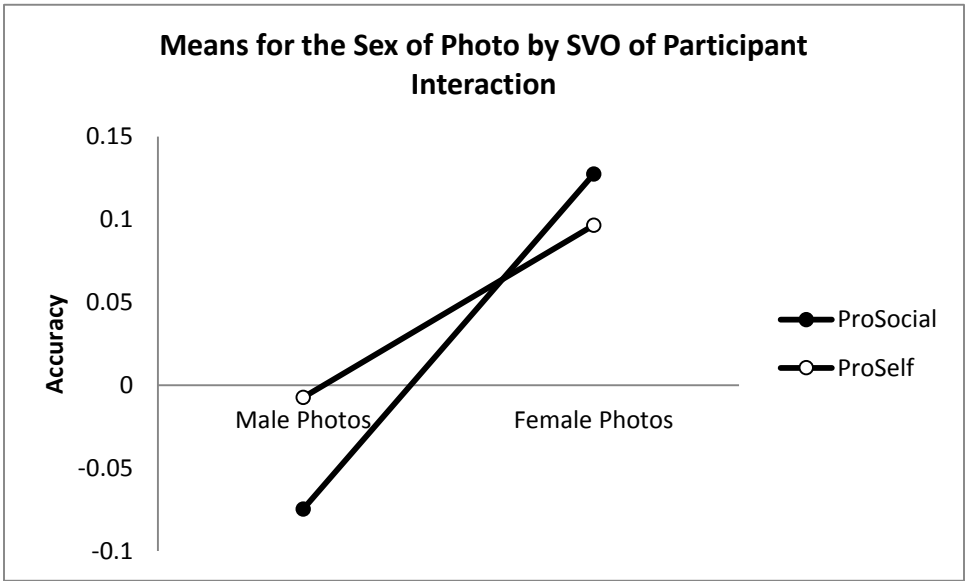

Figure 4. Means for the Sex of Photo by SVO of Participant interaction.

3.3.2.3.2 Effects for SVO of Photo

There was a significant effect ($F(1,141 = 16.27$, $p < 0.001$) for SVO of Photo. ProSocial photos had a higher accuracy score (Mn = 0.339) than ProSelf photos (Mn = -0.259). However, this difference could be in part due to a bias in participants to give higher likelihood ratings of ProSociality. To see if cooperation ratings of ProSocial photos were in fact higher than cooperation ratings of ProSelf targets, the accuracy score for ProSelf targets were multiplied by -1. This modified measure for ProSelfs (Mn = 0.259) was compared with the accuracy score for ProSocials (Mn = 0.339), and a significant difference was observed ($F(1,152) = 4.537$, $p = 0.034$). That the mean for both ProSocial and ProSelfs was positive indicates a bias to give higher likelihood ratings of ProSocialty. That the means are significantly different suggests that ProSocials are seen as more likely to be cooperative than ProSelfs; a form of accuracy. The main effect for SVO of Photo was not moderated by any other variables.

3.3.2.3.3 The Sex of Photo by SVO of Photo Interaction

The means associated with this significant interaction ($F(1,141) = 88.27$, $p < 0.0001$) are shown in Figure 5. The effect for the photo's SVO are much more pronounced in female photos. The means also suggest that a sex-role stereotype (Females are more cooperative than Males) may have been partly guiding the likelihood judgments of cooperation.

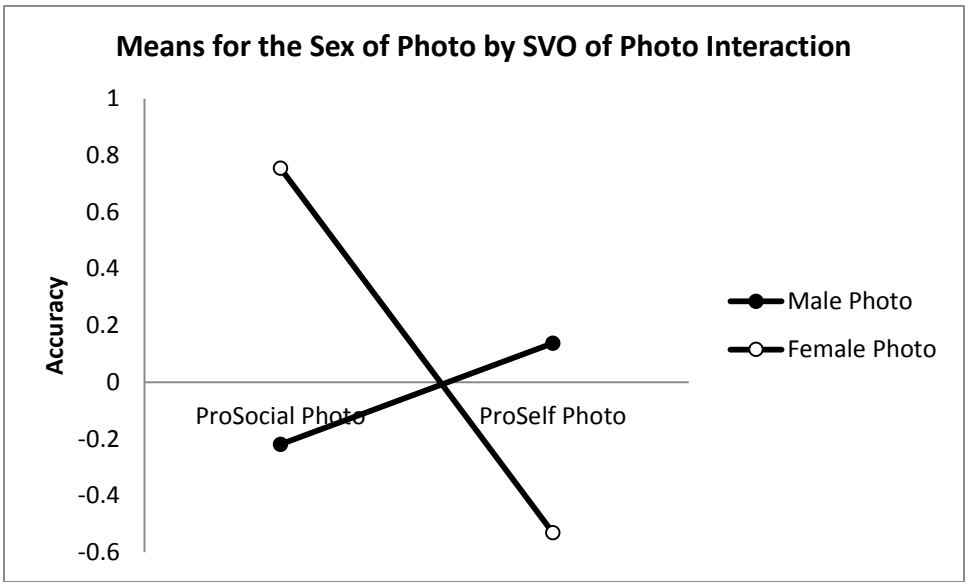

Figure 5. Means for the Sex of Photo by SVO of Photo interaction.

This interaction was not moderated by any other effects. For the same reasons as in the analysis of the effect of SVO of photo the accuracy scores of ProSelf photos (both male and female) were multiplied by minus one. Figure 6 shows the means for these transformed measures. The interaction remains significant ($F(1,141) = 4.53$, $p = 0.034$). Additional tests showed that within Male Photos there was no effect for SVO ($p = 0.270$), but that ProSocial and ProSelf female photos received different cooperation ratings ($F(1,141) = 14.73$, $p < 0.001$). Figure 6 shows a strong bias to rate Females as overall more cooperative than Males. However given that Female ProSocials receive higher ratings than Female Proselfs there is also evidence for some level of accuracy.

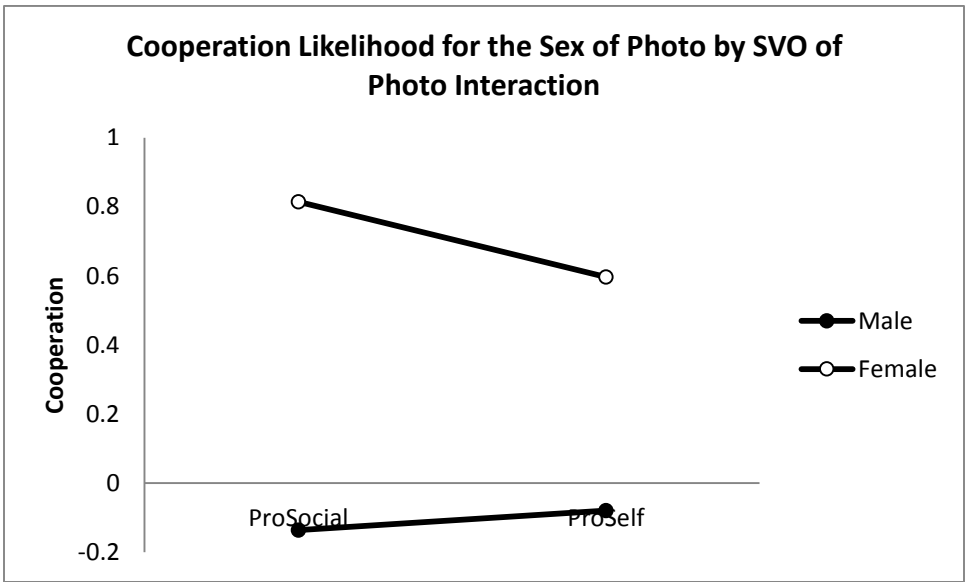

Figure 6. Cooperation Likelihood for the Sex of Photo by SVO of Photo Interaction.

3.3.2.4 Accuracy for Each Sex by Theory of Mind

To check for a differential effect of ToM on Accuracy of Male photos versus Female photos, we ran two more regressions using these transformed Accuracy measures in which (1) Accuracy for Male photos and (2) Accuracy for Female photos were dependent variables predicted by RME score. For Accuracy of Male Photos, the constant is non-significant ($\beta 0 = -0.028$, $p = 0.26$) while RME score is significant ($\beta RME = -0.017$, $p = 0.01$). For Accuracy of Female Photos, the effects are just the opposite: significant constant ($\beta 0 = 0.11$, $p < 0.0001$) but non-significant RME score ($\beta RME = -0.0048$, $p = 0.49$). These results suggest that the SVO of females is detectable, while the SVO of males is not – though judges who score lower on RME can better detect the SVO of males. For judging females, ToM does not appear to play a role.

Chapter 4

DISCUSSION

4.1 General Overview

The present thesis builds upon the work of Shelley et al. (2010) and Page (2012) whose work suggests that individuals are able to detect the Social Value Orientation of an unknown person given a short videotape, or only an emotional photograph. Our results replicate these earlier studies.

The addition to this previous work involved a focus on the possibility of individual differences in Theory of Mind as a predictor of detection accuracy. Specifically, we hypothesized that those who better detected the emotional state of a pair of eyes would be more successful in detecting cooperative / non-cooperative predispositions from a full-faced emotional photograph. Below we discuss our findings for Theory of Mind, which will be followed by discussion of results un-related to Theory of Mind. We conclude with a consideration of the implications of our findings for the evolution of cooperation.

4.2 Theory of Mind

Our results suggest that participants were able to detect SVO above chance, in line with results from Shelley et al. (2010) and Page (2012). Our lack of main effects (for participant's sex, SVO, and viewed emotion) is also consistent with this research. The current project is one of the first to look at Theory of Mind ability and cooperativeness detection accuracy. Unfortunately, we found just the opposite of what was predicted. Theory of Mind scores weakly, but *negatively* predicted accuracy.

This result might not be so unexpected given the findings of another research group. Sylwester et al. (2012) performed two studies: In Study 1, participant were

shown either short (1 to 5 seconds) or long (1 to 2 minutes) video clips of people playing a social dilemma-like game. Participants completed the RME and examined its relation to accuracy in the prediction of the targets' choice in the dilemma game. They found that overall, participants were unable to detect who cooperated and who did not. Participants who viewed the short clips, however, did accurately predict behavior above chance; those who saw the long clips did not. Most interestingly, they found a significant relationship between RME and detection accuracy, but only for those viewing the long clips. Specifically, RME score was positively related to detecting cooperators, but negatively related to detecting non-cooperators. Although this could be interpreted as a bias for rating viewed individuals as more cooperative in general, Sylwester et al. (2012) show that there was no bias for rating either males or females as cooperative in the long clip condition (Sylwester et al., 2012, Supplementary material).

Study 2 in Sylwester et al. (2012) is most relevant to the current thesis. Here, they used the picture stimuli from the previously discussed Verplaetse et al. (2007) study in which photos were taken during the decision-making moment in a laboratory social dilemma. Participants guessed whether or not the viewed person acted cooperatively, and compared their accuracy to two different ToM measures, one of which was the same RME test used in this thesis. They found that participants' estimates of cooperation were significantly *below chance*, and that neither ToM measure correlated with success in cooperativeness detection. Of particular interest to this thesis, a Sex of Photo effect was found for Study 2. Participants were biased in perceiving females as cooperative ($p < 0.01$) and males as non-cooperative ($p < 0.01$)

(Sylwester et al., 2012, Supplementary material). Our results were consistent with this finding (Figure 6).

It must be noted that our results suggested a *negative* relationship between RME and Accuracy from still photos, whereas Sylwester et al. (2012) found *no* relationship between the two. One possible explanation offered by the authors for their results suggests that emotion and motive identification operate via two distinct cognitive mechanisms. If this were the case, our results indicate that these two mechanisms may be at least partially in opposition to each other. This notion, however, fails to explain the positive relationship between Theory of Mind and cooperation detection in the long video clips found in Study 1 of Sylwester et al. (2012). Such a finding suggests that recognizing cooperative predispositions requires more information than do emotions. However, the previous work of Shelley (2010) and Page (2012), replicated in the present thesis demonstrates that the detection of SVO is in fact possible given short video clips, or single still photos. In support of this notion, it is interesting to point out that many individuals are able to identify the emotional state of only a pair of eyes well above chance (Baron-Cohen et al., 2001). In contrast, our observed mean for SVO Accuracy from viewing a full face is small, as are those in Shelley et al. (2010) and Page (2012). Following this idea, the negative relationship between ToM and Accuracy could be explained by interference from the emotion detection system upon the motive detection system.

We found a very small relationship between RME and SVO detection accuracy, but it was the opposite of what we expected. A possible limitation could be our measure of Theory of Mind. It could be asked if (1) the our chosen test is reliable and valid, and (2) if this test was the ideal choice for this thesis.

The RME test is a very popular measure for Theory of Mind (Baron-Cohen et al., 2001; Richell et al, 2003; Paal & Bereczkei, 2007; Declerck & Bogaert, 2008; Hünefeldt et al., 2013), lending superficial support to its reliability and validity. Its internal consistency has been reported in several studies, most finding moderate to acceptable levels of Cronbach's α. (Voracek & Dressler, 2006 found $\alpha = 0.63$ for males, and 0.60 for females; Harkness et al., 2010 found $\alpha = 0.58$; Dehning et al., 2012 found $\alpha = 0.70$; Prevost et al, 2014 found $\alpha = 0.77$). It thus seems that the RME is a reliable measure of Theory of Mind, which was a basis for choosing this test in the present thesis.

In the present thesis, α was 0.729 for male participants, but only 0.360 for females. This raises the possibility that the RME/Accuracy relationship would differ for male and female participants. However, as reported in the Results section, this relationship was not moderated by participant sex ($p = 0.41$). In separate regression analyses (for males and for females) the regression slopes were -0.01 and -0.02 for males and females respectively.

Regarding the validity of the RME, the issue is whether this test correlates with anything. It has been found that RME scores correlate positively with empathy (Declerck & Bogaert, 2008; $r = 0.206$, $p < 0.001$), negatively with age (Henry et al, 2014; $r = -0.43$, $p < 0.001$), positively with heart rate variability (Quintana et al., 2012; $r = -0.29$ $p < 0.05$), and positively with verbal intelligence (Peterson & Miller, 2012; $r = 0.49$, $p < 0.05$) and with combined performance and verbal intelligence (Baker et al., 2014; $r = 0.24$, $p < 0.001$). In addition, a recent study of 353 MTURK workers (Personal communication, Kuhlman) found that the RME was correlated negatively with three personality traits that comprise the so-called "Dark Triad" (2015). For

Psychopathy, $r = -0.39$, $p < 0.001$; for Machiavellianism $r = -0.16$, $p = 0.008$; and for Narcissism $r = -0.298$, $p < 0.001$. It has hard to argue that the results of the present thesis are a result of the general invalidity of the RME. Further, the MTURK results suggest that RME is positively related to measures of psychological adjustment.

We believed that the RME was the best measure of ToM for the current study. Other measures have participants identify an emotional state from a full face (Baron-Cohen, Wheelwright, and Jolliffe, 1997; Ali & Chamorro-Premuzic, 2010) or from a voice (Rutherford, Baron-Cohen, and Wheelwright, 2002; Golan, Baron-Cohen, Hill, and Rutherford, 2007; Ali & Chamorro-Premuzic, 2010). Given that our SVO detection task consisted of participants evaluating whole faces, a measure of ToM using voices seemed less appropriate than one showing at least part of a face. It could be argued that having an individual judge emotional states from whole faces rather than just a pair of eyes would be more appropriate, but we disagree. The ToM Faces task (Baron-Cohen et al., 1997) has participants select from only two emotional states, and has largely been abandoned by its creator in favor of the RME task. The RME task gives four possible emotional states as multiple choice answers. Having more answer choices and showing only eyes rather than a whole face renders the RME task a more rigorous assessment of Theory of Mind. In addition, using the RME test instead of the Faces test helped us avoid redundant tasks in the laboratory (recall that our SVO detection task consisted of rating full faces).

Another limitation could have been the Social Rating task itself. Is it possible that our "choice rules" (Appendix) definition of Social Value Orientations was unclear to participants? As discussed in the Methods section, we gave participants quiz questions to assure that they understand the choice rules. Individuals who missed more

than two questions were excluded from analysis, such that our data comes from participants who understood our explanation of the choice rules. In addition, the choice rules definition of SVO has been used in previous research, producing consistent results (Shelley et al., 2010; Page, 2012). In sum, our results for RME leave us with little reason to believe that this measure has little (if anything) to do with the detection of cooperative intent of strangers.

Before abandoning the idea that Theory of Mind has nothing to do with cooperativeness detection, future research should use different ToM measures (like the ones reviewed above). It would also be important to see if measures other than Theory of Mind can predict SVO detection accuracy. For example, are those who are more empathic better at identifying an individual's cooperativeness? Or perhaps those who perform better on an IQ test, or on a measure of verbal intelligence? It could be that other personality or ability measures could be useful for explaining individual differences in cooperative / non-cooperative predisposition recognition.

4.3 Results Not Related to Theory of Mind

Despite finding a small effect, our results lend evidence to the notion that the cooperativeness of an individual is detectable from rather minimal information, such as an emotional expression.

Research on Social Value Orientation suggests that most individuals have goals of mutual cooperation. (Kuhlman & Marshello, 1975; Balliet, Parks, and Joireman, 2009). Previous work suggests that individuals reveal some information about their social motives nonverbally / implicitly (Schug et al., 2010; Shelley et al., 2010), and that observers are sensitive to this information (Verplaeste et al., 2007; Pradel, Euler, and Fetchenhauer, 2009; Shelley et al., 2010; Page, 2012). We have

provided further evidence that Social Value Orientations are detectable from static facial cues. Our results, in conjunction with previously discussed research, lend support to the optimistic idea that we may not need external coercion to produce outcomes of mutual cooperation.

If cooperative intent is detectable from nonverbal cues, it is possible that cooperative people are able to identify other cooperative people and establish relationships that will result in mutually cooperative outcomes. To this end, we provide evidence that people are sensitive to information about an individual's Social Value Orientation (i.e. their cooperativeness) that is transmitted nonverbally (i.e. conversational mannerisms or emotional expressions). Individuals who can detect cooperative / non-cooperative predispositions of others would be able to form more accurate expectations of those around them, both avoiding exploitation and achieving mutual cooperation at increased rates.

Further, and consistent with earlier research, skill at such detection does not appear to differ as a function of the perceiver's sex or SVO. In general, people appear to be equally good in detecting cooperative intent. Recall the dilemma presented towards the end of the Introduction section: a stranger has asked if he can borrow your cell phone to make one call. Our results suggest that information about how cooperative this stranger's intentions are may be available in the form of nonverbal cues. It is possible that this stranger will implicitly project enough information about himself for us to make an appropriate decision – and not be exploited by a possible thief.

Returning to Goal Expectation Theory, the generality of cooperative intent detection suggests that relationships between two cooperators may more easily and

frequently be formed than those involving at least one non-cooperator. Two cooperators have the goal of mutual cooperation, and are equally good as non-cooperators in detecting the cooperative intent of a stranger. To the degree that these relationship formation processes extend to the formation of mating relationships, we have further insight as to how cooperative motivation may, as many researchers in psychology and behavioral economics argue, has evolved to produce an essentially cooperative species (Bowles & Gintis, 2011).

Finally, the current thesis did not investigate whether participants would *actually* cooperate with individuals perceived as more cooperative; we simply found that participants could correctly identify cooperative and non-cooperative individuals above chance. Although we do not believe it to be a huge leap to suggest that individuals would behave differently with a person perceived to be cooperative versus one perceived to be uncooperative, we did not produce empirical evidence for such a possibility. This is a notable limitation of our work that restricts the extent to which we can generalize our findings to the real world. Regardless, we hold that individuals recognized as cooperative would be *expected* to act as such, which would encourage cooperation from interdependent others. Future work could ask whether or not individuals would cooperate with others perceived to be cooperative, and not cooperate with others perceived to be non-cooperative.

REFERENCES

Ali, F., & Chamorro-Premuzic, T. (2010). Investigating theory of mind deficits in nonclinical psychopathy and machiavellianism. *Personality and Individual Differences, 49*(3), 169-174.

Apperly, I. A. (2012). What is "theory of mind"? concepts, cognitive processes and individual differences. *The Quarterly Journal of Experimental Psychology, 65*(5), 825-839.

Au, W. T., & Kwong, J. Y. Y. (2004). Measurements and effects of social-value orientation in social dilemmas: A review. *Contemporary psychological research on social dilemmas.* (pp. 71-98) Cambridge University Press, New York, NY.

Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution, 54*(1), 39-57.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations, 12*(4), 533.

Balliet, D., Mulder, L. B., & Van Lange, Paul A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin, 137*(4), 594-615.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? evidence from normal adults, and adults with autism or asperger syndrome. *Visual Cognition, 4*(3), 311-331.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241-251.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters,2*(3), 412-414.

Bodden, M. E., Mollenhauer, B., Trenkwalder, C., Cabanel, N., Eggert, K. M., Unger, M. M., . . . Kalbe, E. (2010). Affective and cognitive theory of mind in patients with Parkinson's disease. *Parkinsonism & Related Disorders, 16*(7), 466-470.

Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology, 47*(3), 453-480.

Bosco, F. M., Gabbatore, I., & Tirassa, M. (2014). A broad assessment of theory of mind in adolescence: The complexity of mindreading. *Consciousness and Cognition: An International Journal, 24*, 84-97.

Bowles, S., & Gintis, H. (2011). A cooperative species: Human reciprocity and its evolution. Princeton University Press, Princeton, NJ.

Brown, W. M., Palameta, B., & Moore, C. (2003). Are there nonverbal cues to commitment? an exploratory study using the zero-acquaintance video presentation paradigm. *Evolutionary Psychology, 1*, 42-69.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169-193.

Declerck, C. H., & Bogaert, S. (2008). Social value orientation: Related to empathy and the ability to read the mind in the eyes. *The Journal of Social Psychology, 148*(6), 711-726.

Dehning, S., Girma, E., Gasperi, S., Meyer, S., Tesfaye, M., & Siebeck, M. (2012). Comparative cross-sectional study of empathy among first year and final year medical students in jimma university, ethiopia: Steady state of the heart and opening of the eyes. *BMC Medical Education, 12*, 34.

Duffy, J., & Feltovich, N. (2002). Do actions speak louder than words? an experimental comparison of observation and cheap talk. *Games and Economic Behavior, 39*(1), 1-27.

Ertan, A., Page, T., Putterman, L. (2009). Who to punish? Individual decisions and majority rule in the solution of free rider problems. *European Economic Review, 3*, 495-511.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'reading the mind in the voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 37*(6), 1096-1106.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science, 312*(5770), 108-111.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*(3), 245-256.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859), 1243-1248.

Harkness, K. L., Jacobson, J. A., Duong, D., & Sabbagh, M. A. (2010). Mental state decoding in past major depression: Effect of sad versus happy mood induction. *Cognition and Emotion, 24*(3), 497-513.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28*(3), 826-839.

Hünefeldt, T., Laghi, F., & Ortu, F. (2013). Are anxiously attached women better mindreaders? *Cognitive Processing, 14*(3), 317-321.

Kerr, N. L., & Kaufman-Gilliland, C. (1994). Communication, commitment, and cooperation in social dilemma. *Journal of Personality and Social Psychology, 66*(3), 513-529.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology, 24*, 183-214.

Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology, 46*, 183-207.

Kuhlman, D. M., & Marshello, A. F. (1975). Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology, 32*(5), 922-931.

Kurzban, R. (2001). The social psychophysics of cooperation: Nonverbal communication in a public goods game. *Journal of Nonverbal Behavior, 25*(4), 241-259.

Liebrand, W. B. G. (1984). The effect of social motives, communication and group size on behaviour in an "N-person" multi-stage mixed-motive game. *European Journal of Social Psychology, 14*(3), 239.

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin, 139*(1), 1-34.

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4*(1), 1.

Murphy, R. O., & Ackermann, K. A. (2014). Social value orientation: Theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review, 18*(1), 13-41.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771-781.

Oda, R., Naganawa, T., Yamauchi, S., Yamagata, N., & Matsumoto-Oda, A. (2009). Altruists are trusted based on non-verbal cues. *Biology Letters, 5*(6), 752-754.

Oda, R., Yamagata, N., Yabiku, Y., & Matsumoto-Oda, A. (2009). Altruism can be assessed correctly based on impression. *Human Nature: An Interdisciplinary Biosocial Perspective, 20*(3), 331-341.

Orbell, J. M., Van, d. K., & Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology, 54*(5), 811-819.

Paal, T., & Bereczkei, T. (2007). Adult theory of mind, cooperation, machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences, 43*(3), 541-551.

Page, M. J. (2012). The effect of potential partner social value orientation on attraction to relationships at zero-acquaintance (Order No. AAI3478755). Available from PsycINFO. (1112147451; 2012-99160-112).

Panagopoulos, C. (2014). I've got my eyes on you: Implicit social-pressure cues and prosocial behavior. *Political Psychology, 35*(1), 23-33.

Parks, C. D., Joireman, J., & Van Lange, Paul A. M. (2013). Cooperation, trust, and antagonism: How public goods are promoted. *Psychological Science in the Public Interest, 14*(3), 119-165.

Pradel, J., Euler, H. A., & Fetchenhauer, D. (2009). Spotting altruistic dictator game players and mingling with them: The elective assortation of classmates. *Evolution and Human Behavior, 30*(2), 103-113.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526.

Prevost, M., Carrier, M., Chowne, G., Zelkowitz, P., Joseph, L., & Gold, I. (2014). The reading the mind in the eyes test: Validation of a french version and exploration of cultural variations in a multi-ethnic city. *Cognitive Neuropsychiatry, 19*(3), 189-204.

Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology, 28*, 363-392.

Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., & Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition. *International Journal of Psychophysiology, 86*(2), 168-172.

Richell, R. A., Mitchell, D. G. V., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. R. (2003). Theory of mind and psychopathy: Can psychopathic individuals read the 'language of the eyes'? *Neuropsychologia, 41*(5), 523-526.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders, 32*(3), 189-194.

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior, 31*(2), 87-94.

Shelley, G. P., Page, M., Rives, P., Yeagley, E., & Kuhlman, D. M. (2010). Nonverbal communication and detection of individual differences in social value orientation. *Social decision making: Social dilemmas, social values, and ethical judgments.* (pp. 147-170) Routledge/Taylor & Francis Group, New York, NY.

Sheldon, K. M. (1999). Learning the lessons of tit-for-tat: Even competitors can get the message. *Journal of Personality and Social Psychology, 77*(6), 1245-1253.

Singer, T., Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *The American Economic Review, 95*(2), 340-345.

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. *Child Development, 0*(0), 1-16.

Sylwester, K., Lyons, M., Buchanan, C., Nettle, D., & Roberts, G. (2012). The role of theory of mind in assessing cooperative intentions. *Personality and Individual Differences, 52*(2), 113-117.

Tazelaar, M. J. A., Van Lange, Paul A. M., & Ouwerkerk, J. W. (2004). How to cope with "noise" in social dilemmas: The benefits of communication. *Journal of Personality and Social Psychology, 87*(6), 845-859.

Tella, M. D., Castelli, L., Colonna, F., Fusaro, E., Torta, R., Ardito, R. B., & Adenzato, M. (2015). Theory of mind and emotional functioning in fibromyalgia syndrome: An investigation of the relationship between social cognition and executive function. *PLoS One, 10*(1).

Van Doesum, N. J., Van Lange, Dion A. W., & Van Lange, Paul A. M. (2013). Social mindfulness: Skill and will to navigate the social world. *Journal of Personality and Social Psychology, 105*(1), 86-103.

Van Lange, Paul A. M., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes, 120*(2), 125-141.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel. *Evolution and Human Behavior, 28*(4), 260-271.

Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D:4D) and baron-cohen's "reading the mind in the eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences, 41*(8), 1481-1491.

Wang, Z., & Su, Y. (2013). Age-related differences in the performance of theory of mind in older adults: A dissociation of cognitive and affective components. *Psychology and Aging, 28*(1), 284-291.

Walker, R. F., & Murachver, T. (2012). Representation and theory of mind development. *Developmental Psychology, 48*(2), 509-520.

Xi, C., Zhu, Y., Mu, Y., Chen, B., Dong, B., Cheng, H., . . . Wang, K. (2015). Theory of mind and decision-making processes are impaired in parkinson's disease. *Behavioural Brain Research, 279,* 226-233.

Appendix

SOCIAL RATING TASK INSTRUCTIONS

The social decision task I just told you about is one that's been used here at the University of Delaware for many years, and also by many researchers around the world. In all these studies, it's been found that people tend to make their choices according to one of a number of different choice rules, which we'll explain below. The rules are listed in no particular order; that is, there is no necessary relationship between the order in which we describe the rules and how commonly or frequently they're used by actual people.

Rule 1: Self +, Other +
By this rule, the decision maker chooses the option for which the sum of the points to Self and Other is the largest.

Rule 2: Self +, Other 0
By this rule, the decision maker chooses the option that gives him/her the largest number of points, regardless of the points that the other gets.

Rule 3: Self +, Other -
By this rule, the decision maker chooses the option that produces the largest difference between his/her points and the other's points, or the option that puts him/her the furthest ahead of the other.

Here is an example, for a decision problem with 3 options.

	A	B	C
You Get	70	100	70
Other Gets	70	0	-70

Self +, Other + would choose A.
Self +, Other 0 would choose B.
Self +, Other – would choose C.