

**Sons of Barabbas or Sons of Descartes?
An evolutionary game-theoretic view of
psychopathy**

by

Lloyd D. Balbuena

A thesis submitted to the
Faculty of Graduate Studies and Postdoctoral Affairs
in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy
in Cognitive Science

Carleton University

Ottawa, Ontario

©2010, Lloyd D. Balbuena



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-79631-3
Our file *Notre référence*
ISBN: 978-0-494-79631-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

Psychopathy is characterized by egoism, shallow emotions, impulsivity, and antisocial behavior. Several functional deficits including low fear, a reward-seeking response style, passive avoidance, impaired somatic markers, and language anomalies are associated with psychopathy and an impaired amygdala is implicated as an underlying mechanism. A minority view is that psychopathy is an evolutionarily adaptive life strategy existing at a small frequency in the population. This adaptive view is based on the association between psychopathic traits and reproductive success, and on higher winnings by psychopathic individuals in social games. The goal of this research was to investigate further the adaptive view by means of games played for money and game-theoretic modeling. Forty-eight students belonging to high and low quartiles of 448 males who completed the Levenson Self-Report Psychopathy Scale participated in the Prisoner Dilemma Task (PD), Iowa Gambling Task (IGT) and the Cambridge Gamble Task (CGT). In the Prisoner Dilemma, rate of cooperation varied by *Match Type* (homogeneous vs mixed by *Subject Type*). High psychopathic subjects won more money than low psychopathic subjects in mixed but not in homogeneous matches. In the IGT, both *Subject Types* learned to refrain from the bad decks starting in the second period and picked more consistently from the good decks in the second half compared to the first. In the CGT, a linear mixed model revealed that high psychopathic individuals placed higher bets when the odds of winning were about even. Low psychopathic individuals were better at aligning their bets with the chance of winning, but winnings did not vary significantly between groups. As a model of dynamics in a mixed population, a computer simulation similar to Axelrod's PD tournament was created and the results were consistent with those of human

participants. Overall, the results suggest that psychopathic traits are rational and advantageous in society at large. Hence, psychopathy is likely to exist either as a fraction of the population or a fraction of each individual's behavior.

Acknowledgements

This thesis is dedicated to my mother and my late father (†) who kept our stomachs full and our minds hungry.

Though this document might someday join the company of tomes in a dusty shelf, it is first a product of my six-year instruction at the hands of various mentors whose contribution I now acknowledge. I wish to thank Dr. Andrew Brook and Dr. John Logan for serving as my supervisors. The unfortunate reader endures the agony of reading but once, but those heroic two, went tirelessly through draft after (endless) draft to bring it to its present form. Together with Dr. Jo-anne Lefevre, incumbent head of the Institute of Cognitive Science, they agreed to a long-distance supervisory relationship, one that spans two time zones. I would never have fulfilled my roles as a graduate student and a son without that arrangement.

I also wish to thank Dr. Robert West for sitting in my committee and sharing his knowledge of game theory and sociobiology. A special word of thanks to Dr. Adelle Forth, for carefully reviewing the prospectus, directing me towards important psychopathy literature, and wisely suggesting the use of Carleton's Mass Testing Program. Chris Motz and Nikolina Antonacopoulos graciously assisted me in the mass testing and SONA and Sean Leach administered the experiments.

In my three years in Ottawa, I wish to thank David Leibovitz for many thoughtful discussions and for teaching me how to ice skate in my first winter. Lianne Dubreuil and Colleen Fulton looked out for me beyond the call of official duty. Rob Thomson, Matthew Rutledge-Taylor, and Aryn Pyke shared lots of academic wisdom over pitchers of beer at "Pike's Place."

I am grateful to Dr. Maxwell Pak at Queens University, for teaching me game theory. In Saskatchewan, I thank Dr. Raymond Tempier and Dr. Mansfield Mela for

three productive summers working in the Psychiatry Department and the Regional Psychiatric Centre. I also thank Randy Duncan and Betty Rohr of SHORE and Koroush Larijani for helping me construct a meaningful narrative of the graduate student life.

I thank my present employers, Dr. Alan Casson of the Department of Surgery and Dr. Carl D'Arcy of SKY-RDC for providing me with the wherewithal to bankroll the substantial expenses involved in the project. The statistical and data analytical nature of my jobs came especially handy in data analysis, and the office space was ideal for research.

I thank Conrad and Isabel Aguilera, my foster parents and landlords in Saskatoon. Their late cat Smokey kept me company over countless microwaved meals. Being psychopathic and antisocial in character, Smokey helped keep my focus on this project. Thank you to my friends —Lyman Moreno, Alvin Alvarado, Leila Dominguez, Charie Mosqueda, Ryan Prado, Jam Prado, Cristina Rivera, Roni Duran, VG Dato, Lalaine Mendoza, Ian Necesito, and their families. Many of them served as “guinea pigs” for experimental trials.

I owe a huge debt of gratitude to people I have not personally met but whose bequest to the open source movement gave me very fine thinking tools. They are the authors of R, \LaTeX , and their packages. Special thanks to Dr. Donald Craig at the Memorial University of Newfoundland whose thesis template served as the foundation for many hours of tinkering and customization. Their altruism is beyond Darwinian explanation.

All of what I have written here comes from skills that I did not have before coming to Canada. I thank my previous academic mentors in Ateneo de Manila, the Jesuit priests Fr. Tom Green (†), Fr. Roque Ferriols, and Fr. Mario Francisco. They wasted many hours writing recommendation letters, until a single graduate program

accepted me. The biggest lesson I learned from them is the willingness to always learn something new: the philosophical awareness that all I know is nothing compared with what else remains to be known.

I wish to thank my sister May, my nephew Meshak, and my brother-in-law Mahesh for welcoming me in their house on my many trips from Saskatoon to Ottawa. The hours I spent playing with Meshak's toys remind me that being a student, no matter the age, is really a luxury that allows one to play with ideas. And finally, I thank my girlfriend Lahlah for believing in me and providing motivation to continue in the midst of difficulty.



Table of Contents

Abstract	ii
Acknowledgements	iv
List of Tables	xii
List of Figures	xiv
1 Introduction	1
1.1 Background	1
1.2 Research Questions and Hypotheses	4
1.3 Overview of the Chapters	5
2 Literature Review: The measurement and mechanisms of psychopathy	8
2.1 Origin of the Concept	8
2.2 The Measurement of Psychopathy	11
2.3 Is Psychopathy a Taxon?	14
2.4 Physiological mechanisms underlying psychopathy	17
2.4.1 Poor passive avoidance learning due to low fear	17
2.4.2 The response modulation hypothesis	19

2.4.3	The somatic marker hypothesis	20
2.4.4	Failure to Empathize due to Amygdala Dysfunction	22
2.4.5	Language Anomalies	23
2.4.6	Towards an Integrated View: Mirror Neurons and Psychopathy	25
2.4.7	Relation of MNS with other Theories	28
3	An Alternative View: Psychopathy as An Adaptation	31
3.1	Comment on the Sociobiology wars	31
3.2	Chapter Objective	34
3.3	Psychopathy is subject to evolutionary dynamics	35
3.3.1	Reproductive Success	35
3.3.2	Genetic basis to psychopathy	36
3.4	Psychopathy is adaptive	37
3.4.1	Psychopaths Play a Rationally Competitive Strategy in the Prisoner Dilemma	37
3.4.2	Lying and Deception in Non-Psychopaths	40
3.4.3	Lethal Aggression is Part of our Primate Heritage and Early Human Communities	42
3.5	How Game Theory Fits with the Adaptive View	43
4	Methods	45
4.1	Recruitment of Participants	45
4.2	Description of Instrument	47
4.3	Physical Setting and Experiment Administration	48
4.4	Description of the Games	49
4.4.1	The Prisoner Dilemma	49
4.4.2	The Iowa Gambling Task	52

4.4.3	The Cambridge Gamble Task	53
4.4.4	The Tower of Hanoi	54
4.4.5	Debriefing	55
5	Results	57
5.1	Participant Characteristics and Overall Winnings	57
5.2	Prisoner Dilemma	58
5.2.1	Descriptive Statistics and Time Trend	59
5.2.2	Predictors of Rate of Cooperation and Points Earned	59
5.2.3	Greedy, Generous, and Reciprocal Strategies	64
5.2.4	Section Summary and Comparison to Previous Research	67
5.3	Iowa Gambling Task	68
5.3.1	NET Scores	68
5.3.2	Moving Average of Good Deck Choices	71
5.3.3	Streakiness as a Measure of Learning	73
5.3.3.1	Bad Deck Streakiness	76
5.3.3.2	Good Deck Streakiness	78
5.3.4	Section Summary and Comparison to Previous Research	79
5.4	Cambridge Gamble Task	82
5.4.1	Entire Sample Analysis	82
5.4.2	Restricted Sample Analysis	83
5.4.2.1	CGT Summary Measures	83
5.4.2.2	Modeling Subgroup Betting Patterns	86
5.4.3	Section Summary and Comparison to Previous Research	91
6	An Evolutionary Game Theoretic Model of the Prisoner Dilemma	
	Results	93

6.1	A Virtual Prisoner Dilemma Tournament	95
6.2	Replicator Dynamics of the Virtual PD	98
6.2.1	A Simulation of Replicator Dynamics in the virtual PD	100
6.2.2	Possible Criticisms of the Replicator Dynamic Model	101
6.3	Evolutionarily Stable Strategies	104
6.4	Psychopathy and Evolutionarily Stable Strategies	108
6.5	Psychopathy and Rationality	109
6.6	Chapter Summary	111
7	Discussion	112
7.1	Summary of Experimental Results	113
7.2	Psychopaths and Non-psychopaths: Cut from the Same Cloth	117
7.3	Honesty/Dishonesty Arms Race	119
7.4	The Psychopath's Cognitive Stance	121
7.5	Limitations and Future Research	123
	Bibliography	130
	Appendices	152
A	The Levenson Self-Report Psychopathy Scale	153
B	Physical Set-up	155
C	Prisoner Dilemma Instructions	156
D	Iowa Gambling Task Instructions	158
E	Distribution of 'Cooperate' Moves	160

F	Distribution of PD Outcomes	161
G	Low P Streaks of Bad Deck Choices	163
H	Equilibrium frequencies of High P, Low P	166
I	Equilibrium frequencies of High P, Low P, with varying Initial Con- ditions	168
J	Hamming Distances	169
K	R Code for the CGT LMM	172
L	Evolutionary Scenarios in 2×2 matrix games	173

List of Tables

2.1	Cleckley's Psychopathy Criteria	9
2.2	Factor Structure of the PCL-R	12
4.1	Advertisement in the SONA	46
4.2	Cash Conversion Table	56
5.1	Mean LSRP scores, Age, and 4-Ring Tower of Hanoi Scores (SDs) as a Function of Low vs. High Psychopathy Groups	57
5.2	Descriptive Statistics of Points Won Across Groups for the Prisoner Dilemma, Iowa Gambling Task and Cambridge Gamble Task	58
5.3	Descriptive Statistics by Subject Type \times Match Type for Cooperation and Winnings in the Prisoner Dilemma	59
5.4	Distribution of Strategies	67
5.5	Descriptive Statistics in the IGT Across Groups	69
5.6	Streakiness in Deck Choices	73
5.7	Descriptive Statistics of Bad Deck Streak Lengths Across Groups based on 1,000 simulated Permutations, Mean (SD)	77
5.8	Descriptive Statistics of Bad Deck Losses Across Groups by Half, based on 1,000 simulated Permutations, Mean (SD)	77

5.9	Mean Length (SD) of Good Streaks Across Groups by Half based on 1,000 simulated permutations	79
5.10	Subject Characteristics and CGT Measures	84
5.11	Comparison of Linear Mixed Models for Amount Staked in the CGT	89
6.1	Virtual PD Tournament Winnings Ranking	97
6.2	Mean Winnings in the Virtual PD by Subject \times Match Type	98
6.3	Most ‘typical’ strategies	100
F.1	Self-Advantageous Defection in the Prisoner Dilemma	162
G.1	Low P Strings of Choices	164
G.2	High P Strings of Choices	165
J.1	Hamming distances of High P strategies	170
J.2	Hamming distances of Low P strategies	171

List of Figures

4.1	Prisoner Dilemma, Row Player plays “S”	50
4.2	Prisoner Dilemma, Column Player plays “S”	51
4.3	Cambridge Gambling Task	53
4.4	The Tower of Hanoi task, 4-Rings	55
5.1	Cooperation Time Trend by Match	60
5.2	Mean Rate of Cooperation in the PD by Match Type	62
5.3	Mean Winnings in the PD by Subject and Match Type	63
5.4	Scatterplot of Log(<i>Self-advantageous Defection Index</i>) and <i>Winnings</i>	66
5.5	Net Total Scores in the IGT Between Groups by Period	70
5.6	Moving Average of Good Deck Choices	72
5.7	Moving Average of Wins	72
5.8	Streakiness in the IGT	75
5.9	Bad Deck Streakiness, with Covariates 1 st Half Losses (Centered) = -.0036 , and 2 nd Half Losses (Centered) = -.0218	78
5.10	Individual Subject Trends of Good Choices in the IGT	81
5.11	Money Staked by Box Ratio	87
5.12	95% Predicted Confidence Intervals of Intercept and Slope Variance	90
6.1	Evolutionary Course of High P, Low P Strategies	102

6.2	Evolutionary Course of High P, Low P Strategies, and TFT	107
B.1	Positioning of Participants	155
E.1	'Cooperate' Moves by <i>Subject Type</i>	160
I.1	Different Starting Values of High P, Low P Strategies	168
L.1	Scenarios 1 to 4	174
L.2	Scenarios 5 to 7	175

Chapter 1

Introduction

Empathy is an involuntary breach of individual separateness.

Susanne Langer, philosopher

1.1 Background

Psychopaths are callous, manipulative, dishonest, emotionally detached individuals (Hare, 1998) who often engage in antisocial behavior. Conventional medical wisdom considers psychopathy a disorder to be treated ¹, as it is believed that sooner or later, psychopathic traits will be expressed in destructive antisocial behavior. The harm done to others is believed to be reflexive in the sense that the psychopath is also harmed by ending up in prison, or deserves to be put in prison. Largely due to this prevailing view, most diagnosed psychopaths are found in jail or mental institutions, but there is growing evidence that psychopathy traits exist as a continuum (Guay, Ruscio, Knight, & Hare, 2007; Walters et al., 2007; Murrie et al., 2007; Edens, Marcus, Lilienfeld, & Poythress, 2006; Marcus, John, & Edens, 2004) and not a taxon as traditionally conceptualized (Harris, Rice, & Quinsey, 1994). Psychopathic traits exist in the general population and can be measured within the construct of

¹Many think that psychopathy is resistant to treatment, (See (Harris & Rice, 2006))

normal personality (Widiger & Lynam, 1998; Eysenck, 1987), in the same way that personality traits exist in everyone.

Studies among so-called successful psychopaths (i.e. those able to integrate themselves in society) are limited due to the difficulty in finding *successful* and *properly assessed* psychopaths, i.e. individuals not in jail (a requirement for being “successful”) do not voluntarily submit themselves to an assessment. But the idea of successful psychopaths has been around ever since Cleckley (1941) described psychopaths in the roles of businessmen, playboys, gentlemen, scientists, physicians and psychiatrists. Cleckley called these individuals, “incomplete manifestations” for whom the full-fledged disorder “has not crowded ordinary successful functioning in the outer aspects of work and social relations entirely out of the picture.” (Cleckley, 1941, p. 189)

However, the more interesting question is: *are there functional reasons why psychopathy exists in the population?* An analogous question in economics is, *why are there firms, not just individual buyers and sellers?* It can be argued that individuals can trade bilaterally without forming corporations, and the law of supply and demand will set the price. Coase (1937) argued that firms exist because there are transaction costs in bringing a product to market. For example, in delivering education to students, a school on the one hand, retains teachers instead of having to search time and again. Teachers on the other hand, affiliate with schools instead of looking for students on their own. Hence, greater efficiency is achieved when both parties eliminate the cost of searching for candidates or jobs by binding themselves to each other in an employment contract. Hence, firms exist.

In the same spirit of economic inquiry, it can be asked, “why are there psychopaths?” In my view, the most promising hypothesis is that it is advantageous to be a psychopath. To be a psychopath, in this view, is to be a “genius” of sorts. In

the same way that autistic savants are impaired but still possess phenomenal abilities (e.g. Rain Man), psychopaths have specialized abilities. Psychopaths are “cheaters” (Mealey, 1995; Ellis, 1988; Harpending & Sobus, 1987), shrewd sexual replicators (Seto, Lalumière, & Quinsey, 1995), achievers in an individualistic society (Cooke, 1998) that are somehow well-adapted (Kinner, 2003). To be adaptive in the evolutionary sense, is for there to be a fit between an organism (or a part thereof) to its environment. Since humans and our primate ancestors have lived in groups for thousands of years, a parasite of cooperation probably arose because of advantages to be gained. This project takes the view that to understand psychopaths, it is important to consider the social milieu because manipulation, deceptiveness, and other traits require social living as a context. In my view, finding the right level of explanation for a phenomenon requires an answer to the question: *is this phenomenon meaningful for a Robinson Crusoe way of life?*² Phenomena such as attention or memory as well as some psychopathic traits (e.g. risk-taking, impulsivity) would be relevant to a Robinson Crusoe existence, but I believe that other traits such as lying, manipulativeness, superficial charm would require the presence of a Guy Friday at the minimum. In brief, when the *explanandum* (psychopathy) is social in nature, the *explanans* ought to include the social level of explanation³.

When the subject of inquiry is a group of individuals, whether it is a colony of bacteria, DNA sequences, predators and prey, one is often interested in how the population evolves over time. In this investigation, an important part of “*why are there psychopaths*” is “*will there always be psychopaths?*” Answers to this latter question

²Robinson Crusoe is the main character in Daniel Defoe’s novel of the same name (1791) who is shipwrecked and stranded alone in an island for 28 years.

³Full disclosure: Unlike a strictly gene-centered view of biological Darwinism, a reductionist view that attributes variation in traits to variations in genes, I allow for natural selection to happen at behavioral and cultural levels. I subscribe to Lamb and Jablonska’s *evolution in four dimensions: genetic, epigenetic, behavioral and symbolic* (Jablonska & Lamb, 2005).

invoke the Darwinian concept of adaptiveness—defined as reproductive success (reviewed in Chapter 3). It is not possible to measure the adaptiveness of psychopathy directly, so it will be modeled instead using game theory where the monetary winnings are fitness points. Game theory, according to Dixit and Nalebuff (2008) is the art of strategic decision-making. Decision-making is “strategic” in their sense when it involves anticipating what the opponent is going to do to you. In game theory, what counts as rational is that which maximizes personal gain. Evolutionary game theory weds together evolution and game theory by studying the survival, fixation, or extinction of phenotypes within a population. For evolutionary game theory, the action that maximizes reproductive success is also the rational action. The specific research questions of this investigation are as follows:

1.2 Research Questions and Hypotheses

1. Do psychopaths favor selfish strategies in social dilemmas? Social dilemmas are *social* in the sense that they involve interactions between individuals and *dilemmas* because they involve a choice between a personal vs. a common good. This question will be modeled by the Prisoner’s Dilemma played for real money. Do psychopaths earn more money? *Importance:* If psychopaths earn more money, this serves as evidence for the “parasite of cooperation” view.
2. Do psychopaths fail to learn from punishment? Experimental results among incarcerated psychopaths (reviewed in Chapter 2) reveal behavior that is skewed towards reward and unmindful of punishment. This question will be modeled using the Iowa Gambling Task. *Importance:* It is known that psychopaths are hard to socialize due to low fear. As discussed in pages 17 and following, punishment does not deter psychopaths to the same degree as normals. We

investigate whether this inability to learn from punishment exists in university students with high psychopathic traits.

3. Can risk-taking by psychopaths be advantageous compared to undue caution by non-psychopaths? Put differently, can caution on the part of non-psychopaths sometimes be disadvantageous? This question will be modeled using the Cambridge Gamble Task in which it was shown that too much risk aversion does not maximize returns in investment. *Importance:* If psychopathy is a disorder—which perhaps it is—does it nonetheless have a redeeming economic benefit? If for example, psychopathy affords economic rewards, then there is little hope for its treatment since there is an economic incentive to be sick. Since acting on behalf of one's interests is rational according to economic theory, then in some sense, to be psychopathic is rational.

1.3 Overview of the Chapters

Chapter II is a critical review of literature from psychology and psychiatry. First, I present the origins of the concept of psychopathy and how Hare's Psychopathy Checklist (Hare, 2003) became the preferred measure. Its factor structure is discussed together with alternative formulations. I then address the question whether psychopathy is a discrete category or a continuum, as the answer to this question has implications on my use of university students. Then, I review prominent theories about the physiological mechanisms underlying psychopathy. These theories implicate different causal mechanisms but share the view that psychopathy is a brain disorder. I present a critique of the evidence for each theory, then suggest how current findings regarding the mirror neuron system relate to each theory.

In Chapter III, an alternative view of psychopathy—that it is an adaptation instead of a disorder is discussed. In contrast to Chapter II theories that take an individual brain approach, this view takes an evolving population for an analytical focus. Briefly, this means that the brain is a product of evolution and mutations that contribute to reproductive success tend to be represented in succeeding generations. Since evolutionary psychology is influenced by the acrimonious debates in sociobiology, I devote a section on it to hopefully exorcise its demons from this project. Studies implicating a genetic basis for psychopathy and its adaptive (i.e., advantageous) nature are then reviewed. Because psychopathy is associated with violence and lying, literature coming from the hunter–gatherer lifestyle of our early ancestors is reviewed. Also, evidence coming from the business world in contemporary life in which the behavior of non–psychopaths resembles that of psychopaths is presented. The rationale in doing so is that if something is adaptive, then it ought to be possessed or undertaken by non-psychopaths as well.

In Chapter IV, the method of the present study is described. The Levenson Self-Report Psychopathy Scale (Levenson, Kiehl, & Fitzpatrick, 1995)—the instrument used to measure psychopathy in participants is described. The research questions above are revisited and hypotheses are presented. Importantly, the questions and hypotheses are tested by means of games played for money. The Prisoner Dilemma, Iowa Gambling Task, and Cambridge Gamble Task are each described. The games are chosen because there are theoretical reasons to believe that psychopaths will do well in the Prisoner Dilemma, but not in the Iowa Gambling Task or Cambridge Gamble Task. In all games, the monetary winnings of psychopaths and controls are compared together with measures used in previous investigations.

In Chapter V the findings for the three games are presented after participant characteristics are described. The findings for each game are summarized and then

related to extant literature. Chapter VI is a simulation of the Prisoner Dilemma results in the mold of Axelrod's virtual tournament. The Prisoner Dilemma is chosen because unlike the other two games, it is social in nature and it is in that context that psychopathy's advantageousness comes into play. Chapter VI brings evolutionary game theory framework to bear upon the results, and constitutes the heart of the *psychopathy-as-adaptive* argument.

Chapter VII is a general discussion of the results and its contribution to the literature on psychopathy. The questions posed here in the Introduction are addressed and placed in the context of psychopathy research. Limitations of the analytic framework and the findings are addressed and suggestions for future research are proposed.

Chapter 2

Literature Review: The measurement and mechanisms of psychopathy

*You think i'm no good
But there's nothing i can do
Think again
I know who is fooling who*

Mike Rutherford

2.1 Origin of the Concept

Psychopathy was first reported by Pinel (1818) in his *manie sans delire* concept, describing individuals with intact understanding but impaired affect. The distinction is made between primary and secondary psychopathy, where the former consisted of individuals with shallow, fleeting emotions (low anxiety) as a result of innate deficits by birth and the latter were high anxious ones who engaged in antisocial behavior as a result of poor upbringing and environmental factors (Blackburn, Logan, Donnelly, & Renwick, 2008; Karpman, 1948; Cleckley, 1941). Some researchers (Walker & McCabe, 1973; Rush, 1972) argued that psychopathy is instead a social condemnation, including a suggestion that the term should be dropped in favor of “bastard”

(Cavadino, 1998), implying that the concept is a judgment masquerading as a diagnosis. But the most influential early work on psychopathy is a collection of anonymized cases and a checklist made by the psychiatrist Hervey Cleckley.

Cleckley's *Mask of Sanity* (1941), paved the way for Hare's Psychopathy Checklist (Hare, 1991). Cleckley's criteria for psychopathy are in Table 2.1. Take note of positive health indicators such as #2, #3, and #14. For him the psychopath is a paradoxical individual who manifests no outward signs of disorder (hence wearing a "mask of sanity") but whose emotional blindness renders him incapable of feeling the impact of his actions on others. While the psychopath presents himself normally, all he is doing is successful mimicry and is therefore severely incapacitated while showing all outward signs of normalcy (Cleckley, 1941).

Table 2.1: Cleckley's Psychopathy Criteria

1 Superficial charm and good "intelligence"
2 Absence of delusions and other signs of irrational thinking
3 Absence of "nervousness" or psychoneurotic manifestations
4 Unreliability
5 Untruthfulness and insincerity
6 Lack of remorse or shame
7 Inadequately motivated antisocial behavior
8 Poor judgment and failure to learn by experience
9 Pathologic egocentricity and incapacity for love
10 General poverty in major affective reactions
11 Specific loss of insight
12 Unresponsiveness in general interpersonal relations
13 Fantastic and uninviting behavior with drink and sometimes without
14 Suicide rarely carried out
15 Sex life impersonal, trivial, and poorly integrated
16 Failure to follow any life plan

Source:(Cleckley, 1941, p. 39)

Cleckley is not clear about whether antisocial actions are central to the disorder.

The actions are sometimes done without any goal whatsoever, and not under a strong urge that demands release (Cleckley, 1941). The role of anstisociality continues to be debated today. Hare and colleagues believe that antisocial acts are essential to psychopathy as evidenced by: (a) the predictability of recidivism from Psychopathy Checklist (PCL) ratings (Hemphill, Hare, & Wong, 1998); (b) that childhood antisocial tendencies predict the lifetime stability of psychopathic traits (Frick, Kimonis, Dandreaux, & Farell, 2003); and (c) the homogeneity of the PCL as a scale and its good psychometric properties (Hare et al., 1990; Harpur, Hare, & Hakstian, 1989). A contrary view holds that the antisocial actions should be dropped as these actions are *sequelae*, and not part of the core deficit (Cooke, Michie, & Hart, 2006).

Despite the continuing relevance of Cleckley's concept of psychopathy, the American Psychiatric Association (APA) did not (and still does not) use the term and instead uses "antisocial personality" (APD) in successive versions. The reluctance of the APA to use Cleckley's terminology lies in its uneasiness with a trait-based description, preferring behavioral indices instead. Cleckley's view is closest to the Diagnostic and Statistical Manual-II's (American Psychiatric Association, 1968) formulation of an antisocial personality (Lilienfeld, 1994). However, the Diagnostic and Statistical Manual-III (American Psychiatric Association, 1980) totally broke away from the DSM-II formulation, having no items in common. The proposed Diagnostic and Statistical Manual-V, planned for release in May 2013, finally restores the classification "psychopathic/antisocial" among the personality types, with type and trait ratings for assessment (American Psychiatric Association, n.d.).

The DSM has been criticized for several shortcomings. Some of them are: (1) The APD list of items yields 397,683 possible variations of APD and therefore the chances of a unitary construct being tapped are low (Cunningham & Reidy, 1998). (2) There is no cut-off score of APD that defines a threshold at which somebody is

likely to be a psychopath (Cunningham & Reidy, 1998). (3) APD includes a high proportion (at least 50 percent) of inmates whereas the same proportion identified by Hare's PCL is only at 13 percent (Widiger et al., 1996). In forensic settings, most individuals meet the criteria for APD but most of them are not psychopaths. By contrast, most offenders with high PCL scores also meet the APD criteria (Hare & Neumann, 2006). It remains to be seen if the soon-to-be-released DSM-V will meet greater acceptance among mental health practitioners.

In the next section, we describe how Cleckley's seminal ideas have been improved upon and incorporated in the Hare PCL-R. The debate regarding the centrality of antisocial actions also rages on.

2.2 The Measurement of Psychopathy

The gold standard for the measurement of psychopathy is Hare's Psychopathy Checklist-Revised (PCL-R) which is based on Cleckley's list of 16 traits (Hare & Neumann, 2006). Using file reviews and an interview, a trained clinician gives a score of 0 (definitely not present), 1 (uncertain), or 2 (definitely present) to each item (Table 2.2). The ratings are based on an interview by a trained clinician and on file reviews. A cut score of 30 in North America and 25 in Europe is generally accepted as a threshold.

PCL was put together using factor analysis, a technique that aims to discover a latent construct hypothesized to account for the variance in a measure, such as test scores (Neumann et al., 2007). Factor analysis follows a logic similar as to linear regression in which the variance is partitioned into a systematic part and a random part. Two major steps are undertaken in factor analysis: (1) the extraction of factors and (2) rotation of the factors to facilitate interpretation (Child, 2006). Two kinds

Table 2.2: Factor Structure of the PCL-R

Factor 1: Interpersonal

Glibness/superficial charm
Grandiose sense of self-worth
Pathological lying
Cunning/manipulative

Factor 2: Affective

Lack of remorse or guilt
Shallow affect
Callous/lack of empathy
Failure to accept responsibility for own actions

Factor 3: Lifestyle

Stimulation seeking
Impulsivity
Irresponsible
Parasitic orientation
Lack of realistic goals

Factor 4: Antisocial

Poor behavior controls
Early behavior problems
Juvenile delinquency
Revocation of conditional release
Criminal versatility

(Source:Hare & Neumann, 2006, p. 77)

of rotation are possible: orthogonal, when the factors do not correlate, and oblique when they do.

These factors are interpreted as a latent construct responsible for the correlation of scores (or measures). The correlation between the original test items and each factor is called the factor loading. For example, in the original two-factor structure of the PCL, some items loaded onto the interpersonal— affective dimension and the others, onto the impulsive—antisocial dimension. In effect, the two dimensions (or factors) group the items to account for the explained variance (called communality). By explaining psychopathy in terms of the factors instead of the original items, one achieves economy and generality. Generality is gained because different measurement instruments (e.g. different versions of the PCL-R) have different numbers of items but should tap into the same underlying factors.

In reality, factor analysis is not a single procedure but has variants such as exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). In EFA, the objective is to infer a model from a given set of data while in CFA, one specifies the model beforehand and tests how well it fits the data using various measures (e.g. goodness of fit index, normed fit index, root means squared residual) (Bryant & Yarnold, 1995). EFA has been extensively used to infer the dimensions of intelligence hypothesized to be responsible for test scores (Guilford, 1956; French, 1951). Because in CFA, one selects the best among competing models, the researcher chooses the “golden mean” between a null model (with no factors and no factor loadings) and a fully saturated model that explains all the variance (Bryant & Yarnold, 1995) but which might not generalize to the population in question.

Several factor analytic studies have been conducted on the PCL/PCL-R. The well-known two-factor solution was found in which *F1* was called the selfish, callous, remorseless use of others and *F2* called chronic antisocial behavior which are corre-

lated at 0.5. (Hare et al., 1990; Harpur et al., 1989). Validation of the two-factor solution via CFA gave poor to moderate fit (McDermott et al., 2000; Darke et al., 1998; Brandt et al., 1997). Cooke and Michie (2001) produced a three-factor solution using a subset of 13 PCL items, excluding the ones pertaining to antisocial behavior. This three-factor model consisting of *F1*: Arrogant and Deceitful Interpersonal Style, *F2*: Deficient Affective Experience, and *F3*: Irresponsible Lifestyle provided superior statistical fit compared to the two-factor model.

However, because of the excluded items, this three factor solution is criticized by Hare and colleagues as being arbitrary and subjective (Hare & Neumann, 2006). The three-factor model also relies on testlets¹, without which, the model had a poor fit in one replication study (Vitacco, Neumann, & Jackson, 2005). A competing model by Neumann (2006) found a four-factor solution to the PCL-R using a North American sample of 5,964 adult offenders (Hare & Neumann, 2006). Two variants of this four-factor solution are: (a) using the individual items to load directly onto the factors and (b) collapsing the items into eight parcels, with two parcels loading onto each factor. A similar four-factor parcel solution was found for adult and adolescent offenders (Hare & Neumann, 2006; Neumann et al., 2006; Forth et al., 2003). See Table 2.2. The four-factor model provides a better statistical fit to the two-factor model (Neumann et al., 2007). In addition, it achieves better statistical fit without the use of testlets and without dropping any of the original list of items (Vitacco et al., 2005).

2.3 Is Psychopathy a Taxon?

Before presenting the results of previous studies regarding the taxonicity or dimensionality of psychopathy, it is important to define what is meant by a taxon, what

¹A testlet is defined as “an aggregation of items based on a single theme” (Ip, 2010)

procedures are used to determine taxa, and why it matters. According to Meehl (2004), a taxon is a type or kind that has more or less “clearly demarcated properties” (p. 39). Usually, something is a taxon if an entity belonging to its domain is either a member or not a member. For example, in the domain of eating utensils, something is either a spoon or not a spoon, a fork or not a fork. There are no intermediates such as a spork. The set of procedures used to determine taxonicity is known as taxometrics. Taxometrics involves identifying taxa and assigning individuals as belonging or not belonging to those taxa (Meehl, 1992). Why does it make a difference that something such as psychopathy is taxonic or dimensional? The crucial difference is that if psychopathy is a taxon, the task of the psychologist is to identify the best cut-off score for a measure, whereas if it is a dimension, one would like the measuring instrument to disperse individuals throughout the continuum.

The family of statistical procedures developed by Meehl for the detection of a taxon relies upon the consistency of outcomes i.e. “convergence of evidence” from multiple procedures. These procedures include MAXCOV (maximum covariance) and MAMBAC (mean above minus below a cut), MAXEIG (maximum eigenvalue) and L-Mode (latent mode factor analysis) (Meehl, 1995). When the latent construct (e.g. psychopathy) is a taxon, the diagnostic factors will produce graph patterns that are distinguishable from those when the latent structure is a continuum (D. Williamson et al., 2002).

In one of the early taxometric studies of the PCL-R, Harris et al. (1994) found evidence for a taxonic latent construct especially for Factor 2 items and variables related to childhood history. Graphs produced by Meehl’s methods (MAMBAC and MAXEIG) showed an inverted U-shape indicative of a taxon, as opposed to the absence thereof. The stronger evidence for Factor 2 taxonicity would be confirmed in two subsequent studies that used antisocial personality disorder items from DSM-IV

(Skilling, Harris, Rice, & Quinsey, 2002; Skilling, Quinsey, & Craig, 2001). Importantly, Skilling et al. did not find taxonicity for personality based characteristics. In another study (Marcus et al., 2004) used the PPI to detect taxonicity and concluded that psychopathic personality has a dimensional latent structure. The graphs produced by MAMBAC and MAXEIG did not produce the inverted U-shape indicative of taxonicity. Marcus et al. argued that if psychopathy is dimensional, this should remove the stigmatizing label of “psychopath” as it would simply be shorthand for “having high psychopathic traits”. It might be argued that the opposite conclusions of the PPI-based vs PCL-R based studies are due to the use of different instruments. However, subsequent studies that used the PCL-R yielded results consistent with a dimensional latent structure. Addressing some criticisms of the Harris et al. (1994) study, Edens et al., (2006) included interviews and used the three-point scoring method as the standard PCL-R assessment should be. In addition to the absence of an inverted U-shape, the L-mode analysis also produced a unimodal as opposed to a bimodal distribution characteristic of taxonicity. Guay, Ruscio, Knight and Hare (2007) undertook the largest taxometric study on the PCL-R with a sample of 4,865 inmates and reported that both the components of psychopathy and their totality have a latent dimensional structure. Walters, Duncan, and Mitchell-Perez (2007) reached the same conclusion with a smaller ($n = 409$) but more heterogeneous sample in terms of inmate security level (maximum, medium, and minimum).

In summary, nearly all researchers agree that the PCL-R is the best measurement tool but there is continuing debate about what should and should not be included in the construct of psychopathy. Factor analysis is an important development in studying constructs that are believed to underlie directly measurements. Like other statistical tools however, its use is a mix of science and art because certain matters, such as how many factors should be extracted and what kind of rotation is to be

performed, depend on the judgment of the theorist. Often this judgment depends on the theory that one has in mind before doing the analysis (Gorsuch, 1990). The search for factor analytical solutions of the PCL and PCL-R is but another arena where the debate about the importance of antisocial actions in psychopathy takes place. On the issue of taxonicity, more recent and better-designed studies have concluded that the underlying structure of psychopathy is dimensional rather than taxonic. This finding has important implications because in the legal system the label “psychopath” can often be used as a basis for indeterminate jail time (Cunningham & Reidy, 2002). For the purpose of the present study, the dimensional nature of psychopathy also justifies psychopathy research in college students. In the next section, the physiological and functional mechanisms underlying psychopathy are discussed

2.4 Physiological mechanisms underlying psychopathy

In this section, leading theories explaining the functional and physiological bases for the traits and behaviors of psychopaths are reviewed. The common denominator of these theories is they take an *individual brain* approach to psychopathy. Typical of each theory’s approach is finding a causal mechanism in the brains of psychopaths to explain psychopathic traits. Chapter 6 brings in a population perspective—a necessary perspective because psychopathic traits cannot be fully deployed in a Robinson Crusoe existence.

2.4.1 Poor passive avoidance learning due to low fear

In experimental psychology, active avoidance learning is exemplified by a rat’s escaping to a safe chamber when a light signaling a forthcoming shock comes on, i.e. commission of an action. Passive avoidance learning, by contrast, is exemplified by

refraining from an action that is associated with punishment (i.e., omission), where such omission is not the main goal. By this distinction, passive avoidance is involved when a rat is being trained to extinguish a learned response by omitting a reward.

Lykken (1957) proposed the low fear hypothesis which holds that psychopaths have intact active avoidance (subserved by the Behavioral Activation System) which allows them to learn from reward, but an impaired passive avoidance system (subserved by the Behavioral Inhibition System), making them poor at learning from punishment.² In a mental maze task, psychopaths were as good as controls in learning correct choices but poorer at avoiding shocked choices. This dichotomy reflects a bias for seeking rewards without fear of punishment. In real life, this dichotomy could be expressed in killing a crime witness without fear of probable imprisonment (Fowles & Dindo, 2006). The low fear hypothesis supports Cleckley's conjecture that psychopaths cannot be socialized as a result of deficient fear conditioning.

Early studies such as those described above, relied on autonomic measures (i.e., electrodermal activation (EDA) and heart rate) as indices of fear and arousal. Patrick (1994) criticized these measures as not being specific and possibly confounded. He introduced the startle blink reflex as a more refined measure of fear. For non-psychopaths, the startle reflex is boosted by aversive stimuli but dampened by pleasant ones. For psychopaths, the dampening effect was present with pleasant stimuli, but the reflex did not grow stronger with unpleasant images (Patrick, 1994). The magnitude of the startle reflex is inversely correlated to high scores in Factor 1 (Vanman et al., 2003; Benning et al., 2005) but positively related to Factor 2 (Vanman et al., 2003). These findings suggest that it is the personality and affective components of psychopathy that relate to low fear. Interestingly, individuals with borderline per-

²BAS activates approaching rewards, BIS applies the brakes to avoid punishment. (Gray & McNaughton, 2003)

sonality disorder, who suffer from intense anxiety, do not have enhanced startle blink reflexes (Herpertz & Koetting, 2005). The startle reflex deficit in emotionally detached psychopaths (Factor 1) is present both for those in jail and in the community (Benning et al., 2005; Justus & Finn, 2007). Women at high levels of psychopathy did not have an impaired startle blink reflex (Justus & Finn, 2007). In terms of the neural basis of low fear, Birbaumer (2005) reported that non-psychopathic subjects (but not psychopathic ones) had increased brain activation in the limbic prefrontal circuit (i.e. amygdala, orbitofrontal cortex, insula, and anterior cingulate) when neutral faces were paired with an aversive stimulus.

Critique. Experimental evidence supports some form of diminished fear in psychopaths. What is most problematic is the lack of consensus on a definition of fear. Rock climbers, decorated heroes, and drug users engaging in risky behavior have different kinds of diminished fear (Levenson, 1990). Is fear a general outlook in life as suggested by Cleckley originally, or is it situational such as in anticipation of pain, or relative to an antisocial lifestyle as suggested by Frick (1995)? Does low fear equal low anxiety? A nonsignificant but suggestive result showed a positive correlation between psychopathy and anxiety, but a negative one between psychopathy and fear (Schmitt & Newman, 1999). Importantly, the PCL-R does not contain an item that indexes fear so either it leaves out an important variable or fear is highly correlated with other measures and it is necessary to determine which one is more fundamental.

2.4.2 The response modulation hypothesis

Newman (1998) suggests that psychopaths have a cognitive (i.e. information-processing) deficit that works to their detriment in reward-based experiments. This deficit is evident when initially, a dominant response is built and then later extinguished. In a card-playing experiment, once a dominant response is rewarded, it is

hard to extinguish this response with punishment (Newman et al., 1987). The experiment also revealed that when cumulative feedback was given (of how much has been won or lost so far) and psychopaths were forced to pause for five seconds before making a move, perseveration decreased.

In another experiment (Newman & Kosson, 1986), psychopaths did a Go/No-go task under two conditions: (1) Reward and Punishment (R + P) and (2) Punishment only (P). In the R + P condition, participants could either win or lose money while in the P condition, they could only lose money. Psychopaths had more errors of commission compared to controls in the R + P condition but not in the P condition. The result suggested that psychopaths fail to control reward-seeking behavior at the cost of receiving more punishment. Cognitive overload is hypothesized to be the root of insensitivity to punishment, where this cognitive overload is defined as “effortful re-allocation of attention” (Newman, 1998, p. 95).

Critique. Both the low fear and the response modulation hypotheses use for a basis experimental results implicating passive avoidance deficits. Where the latter differs from the former is that cognitive variables such as attention and information-processing are identified as the cause (Lorenz & Newman, 2002) and passive avoidance is secondary. Their finding that a five-second delay decreased perseveration is open to an alternative interpretation: impulsivity leads to passive avoidance errors. Thus, emotional factors (other than fear) might also be present in psychopaths as will be discussed in the next section.

2.4.3 The somatic marker hypothesis

In this view, psychopaths engage in risky behavior because they lack the bodily feedback that galvanizes one to act upon what is known. According to the somatic marker hypothesis, decisions are made well before they reach consciousness through

bodily feedback that steers the individual towards advantageous choices (Damasio, 1994). The hypothesis found experimental validation in the Iowa Gambling Task (Bechara et al., 1997). In this card game, participants draw one card at a time from 4 decks, two of which are “good”, leading to net wins and the two are “bad”, leading to net losses. Similar to the response modulation paradigm, the bad decks are conditioned for reward initially but gradually become aversive in the form of huge losses. In a typical experimental set-up, the performance of normal controls is compared with impaired individuals, in this case those with ventromedial prefrontal cortex (VPC) damage but intact intelligence and memory. In the IGT, impaired individuals, unlike normal controls, continue to make bad choices even when they can state verbally which decks are good or bad. Unlike normal participants, impaired individuals do not develop anticipatory skin conductance responses (SCRs) before selecting from the bad decks. According to somatic marker theory, the absence of SCRs reflects a lack of bodily feedback representing previous emotional experience from similarly risky events. Although Damasio and colleagues did not use psychopaths in their experiment, it is suggested that both psychopaths and VPC-impaired individuals share a common failure to learn from punishment (Damasio, 1994, 2000).

Four studies have investigated the performance of psychopaths in the IGT. Schmitt, Brinkley, and Newman (1999) failed to find a relationship between psychopathy and risk-taking, but found a link between anxiety and risk-taking. Blair, Colledge, and Mitchell (2001) found that psychopathic boys (measured by the Psychopathy Screening Device (Frick & Hare, 2001)) aged 9 to 17 years were more likely to make risky choices compared to controls. A third study among Dutch students found that psychopathic students continued to draw from the bad decks unlike non-psychopaths (Honk, Hermans, Putman, Montagne, & Schutter, 2002). This last study was limited by the lack of a psychopathy instrument as such. Psychopathy was instead estimated

from combined Behavioral Inhibition System / Behavioral Activation System scores. Finally, Mitchell et al. (2002) found a sustained tendency by psychopaths to select high-risk cards compared with controls in an incarcerated sample.

Critique. In summary, three out of four studies found a similar high-risk pattern of card choices in psychopaths. Two criticisms of the somatic marker hypothesis are as follows: (1) Low activation in the prefrontal–limbic areas among psychopaths might mean higher activation in brain areas sensitive to reward. If so, psychopathy would not be a matter of total absence of somatic markers but of impaired *fear* somatic markers. For example, a recent study among substance disorder individuals revealed a group that was insensitive to any future outcome (reward or punishment) and a group that was hypersensitive to reward but insensitive to punishment (Bechara et al., 2002). (2) For the somatic marker hypothesis to explain psychopathy, there needs to be a task that is social in nature. If psychopathy is related to manipulateness, egocentricity, lack of responsibility and the like, it would be important to study their behavior in a context where interaction with another player is required. This way, we can tell whether disadvantageous personal choices are likewise disadvantageous in an interpersonal context. The present investigation takes this direction.

2.4.4 Failure to Empathize due to Amygdala Dysfunction

Like the somatic marker hypothesis, Blair’s so-called Violence Inhibition model or Integrated Emotion System (Blair, Jones, Clark, & Smith, 1997; Blair, 1995) model implicates the dysfunction of the ventromedial prefrontal cortex (VPC). However, Blair suggests that the VPC fails because it receives reduced inputs from the amygdala (Blair, 2007). In his view, the amygdala is what gives goodness or badness to objects, actions, and individuals. With a dysfunctional amygdala, psychopaths do not empathize when faced with distress cues, such as when a prospective victim screams

out of fright. In effect, the psychopath's knowledge that striking somebody is wrong lacks the power to inhibit a stronger impulse to harm the victim.

Critique. There is evidence that activation of the amygdala during a task involving emotion is reduced in psychopathic individuals (Gordon, Baird, & End, 2004; Kiehl et al., 2001). Rilling et al. (2007) found that high psychopathic participants in a Prisoner Dilemma game had lower activation in the amygdala compared to healthy participants when experiencing an aversive outcome. However, other studies point the neural dysfunction in psychopathy to other areas such as the orbitofrontal area (Veit et al., 2002) or the prefrontal cortex (Raine, Lencz, Bihrlé, LaCasse, & Colletti, 2000). Overall, the dysfunction does not appear to be exclusive to the amygdala but includes prefrontal and hippocampal areas. So, Blair is probably correct that the amygdala is dysfunctional in psychopaths but that it is important to consider circuits with bi-directional feedback loops. It is also possible that the mechanisms identified by these researchers subserve the functional deficits of poor passive avoidance and response reversal problems mentioned earlier.

2.4.5 Language Anomalies

The link between psychopathy and a lack of ability to understand linguistic meaning was suggested first by Cleckley (1941) who portrayed psychopaths as having semantic aphasia. These aphasics are capable of producing language, but the words and sentences they produce have no meaning. Cleckley argued by analogy to semantic aphasia that psychopaths do not really know the meaning of emotions, but simply pick up cues about meaning from people's behavior. A related analogy is that of color blindness. Psychopaths are blind to emotion in the same way that colorblind individuals cannot see colors correctly, although they might infer facts about apples from living with other people.

It is impossible for him to take even a slight interest in the tragedy or joy or the striving of humanity as presented in serious literature or art. He is also indifferent to all these matters in life itself. Beauty and ugliness, except in a very superficial sense, goodness, evil, love, horror, and humor have no actual meaning, no power to move him. (Cleckley, 1941, p. 40)

It might be argued that language impairment is a manifestation of an underlying emotional impairment (e.g. fear or empathy). Laboratory experiments, however, have demonstrated the independence of language from emotional deficits. Kiehl et al. (2004) found that in affectively neutral words, psychopaths are significantly slower than control participants in processing either abstract or concrete words. This study used fMRI to determine the areas of the brain involved in processing concrete or abstract words in both controls and psychopaths. Consistent with the result of an earlier study (Kiehl et al., 1999), the control participants showed greater activation of the right hemisphere compared to psychopaths in processing abstract vis-à-vis concrete words. This result, along with prevailing literature (Beauregard et al., 1997; Bottini et al., 1994; Brownell et al., 1983) suggests that the right hemisphere is favored in metaphorical, humor, or “image-based” meanings. Psychopathic participants, on the other hand, did not exhibit greater activation of the right hemisphere in general and the superior temporal gyrus in particular (Kiehl et al., 2004). Since only affectively neutral words were used in these experiments, the conclusion is that psychopaths’ impaired ability to experience guilt, empathy, or remorse might be a manifestation of an impaired cognitive ability to think in the abstract (Kiehl et al., 2004).

Other language impairments in psychopaths that have been suggested are: weaker left hemisphere activation (Hare & Jutai, 1988); lack of *emotion facilitation* in word identification (Lorenz & Newman, 2002; Intrator et al., 1997); having hand gestures that were unrelated to speech content (Gillstrom & Hare, 1988); and incoherent speech when describing emotional or personal events (S. Williamson, Harpur, & Hare, 1991).

Critique. As mentioned, the theorists in this group are circumspect in their claims. If it were claimed that psychopathy has its origins in a language disorder then an account would have to be given why individuals with Tourette's syndrome (especially the subpopulation given to uttering vulgar language) do not also have the antisocial tendencies and personality traits of psychopaths. The language deficits are raised here because they do serve as additional evidence for the different way that the brains of psychopaths function.

Thus far, I have reviewed *a-personal* causal mechanisms, i.e.—low fear and passive avoidance—whose stimuli need not be people, then moved on to more properly *personal* mechanisms, such as violence inhibition and language anomalies. The next theory looks at brain function with an explicit social dimension, taking the view that specific brain functions arose to navigate the social milieu.

2.4.6 Towards an Integrated View: Mirror Neurons and Psychopathy

In this section, I review mirror neuron system (MNS) theory and argue its merits in possibly unifying the findings of the various theories given previously. Because the theory is so new, only one study has been published. The hypothesis is very straightforward and readily testable: psychopath traits and behaviors (especially lack of empathy) are due to a dysfunctional mirror neuron system (Keysers & Gazzola, 2009). The single experimental result (thus far), does not support the hypothesis, and needs verification and is described starting in page 27.

The key discovery in mirror system theory is that the same neural system is involved whether doing an action oneself or observing it in another individual (Dapretto et al., 2006). Mirror neurons were first discovered in the premotor area (F5) of macaque monkeys (Gallese et al., 1996). In neuroscience, brain systems are tradi-

tionally divided into the sensory system and the motor system. Similarly neural signals are classified either as afferent (incoming) or efferent (outgoing). This traditional terminology suggests a neat dichotomy between “perceiving” and “acting” with the brain’s central executive linking the two. But the firing of mirror neurons in the motor area during observation indicates an overlap of the neural systems involved in perception and in action. This overlap has huge implications in terms of psychopathy because the capacity to empathize³ is impaired if psychopaths have a dysfunctional MNS. If so, this finding would minimize psychopaths’ culpability. Just as a blind person can be excused for bumping onto people on the street, then psychopaths—if they truly could feel no emotional pain—would be less liable for the harm they inflict.

Mirror neurons have cognitive and affective roles in social interaction. While the two aspects are related, the cognitive part has to do with perspective-taking while the affective part is a matter of sharing the sensory and emotional states of the other person (Hein & Singer, 2008). Each of these aspects is described next.

Cognitive. Mirror neurons code for the intention behind another individual’s action (Iacoboni & Dapretto, 2006). This conclusion is supported by several lines of evidence: (1) Unlike “grasping neurons” that fire at the sight of a graspable object, mirror neurons do not fire automatically. (2) The level of firing of mirror neurons varies depending on the salience of the action. They do not fire when the monkey sees a pantomimed grasping action but where there is no object to be grasped (Gallese et al., 1996). (3) They fire when an observed action is logically related to and precedes a goal, e.g. grasping a piece of food and bringing it to the mouth (Pellegrino et al., 1992). Taken together, these findings suggest that mirror neurons represent what is in the mind of an observed party while doing an action.

³The argument here is that the intensity of one’s own mirror neuron activity correlates with one’s ability to “feel into” (German: *einfühlung*) another person’s experience

Affective. The link between mirror neurons and empathy is elucidated in studies comparing autistic with typically developing children. Although both groups are capable of imitation facial expressions, autistic children lacked mirror neuron system (MNS) activity when imitating faces showing anger, fear, happiness, neutral, or sad expressions (Dapretto et al., 2006) as revealed by fMRI results. In typical children there was activation in the *pars opecularis* together with limbic structures (i.e., centers of emotion) and none for in the autistic children. More telling is the finding that among the autistic children, the activity of the MNS was inversely proportional to their social skills. Instead of MNS activity during imitation, the autistic children had heightened activity in different brain areas (i.e., anterior parietal and right visual association areas). Although not strictly related to psychopathy, this finding in autistic children suggests an avenue for research since like them, psychopaths are known to have shallow affect, and might perhaps suffer from the same underlying deficits.

In the only published study so far on MNS and psychopathy, college students with the highest coldheartedness score in the Psychopathic Personality Inventory PPI, (Lilienfeld & Widows, 2005) had the greatest reduction in motor evoked potential⁴ when viewing a needle penetrating a human hand (Fecteau, Pascual-Leone, & Théoret, 2008). In this experimental paradigm, this finding means that, counter-intuitively, the most coldhearted individuals are the most responsive to pain in others. It is suggested that the experiment may have tapped into sensory empathy and not emotional or trait empathy. The greatest limitation in the study is the use of a healthy sample of college students and a low sample size of 18 individuals. More replications are needed to investigate the link between MNS and psychopathy.

⁴MNS theories work with the observation that the impact of observing a painful experience of another person is proportional to the activity of one's mirror neurons. The greater the motor evoked potentials in transcranial magnetic stimulation (TMS), the greater one "felt" the other's pain. To put this principle very simplistically, I am hurt by seeing another person in pain.

A word of caution needs to be said in assuming that autism and psychopathy are similar in functional or neurological performance. While both autism and psychopathy are social dysfunction disorders, their deficits in empathy are not equivalent (Blair, 2005). Empathy is not a unitary system but can be fractionated into (a) cognitive empathy (*Theory of Mind or ToM*); (b) motor empathy which allows for imitation and is subserved by the mirror–neuron system; and (c) emotional empathy which involves having the appropriate emotional response to stimuli. Furthermore, it is not the case that these types are equally impaired in psychopaths and autistic individuals. Psychopaths are not impaired in ToM (Blair, 2008; Blair et al., 1996) and are selectively impaired in sadness, fear, and disgust (Kosson et al., 2002). Autistic individuals seem to be impaired in all three forms of empathy. They have ToM deficits (Baron-Cohen, Leslie, Frith, et al., 1985) and motor empathy deficits as shown by the absence of MNS activity. This absence of MNS activity in turn indicates an absence of the “inward” emotional component of the facial expressions of emotion, echoing Cleckley’s observations.

2.4.7 Relation of MNS with other Theories

How might a defective MNS system mesh with previous research findings? The MNS is the basis of imitative learning (Iacoboni & Dapretto, 2006). So if psychopaths have a dysfunctional MNS, their poor learning in some situations would follow. For example, the low fear of psychopaths (as evidenced by Patrick’s startle blink reflex) can be attributed to a dysfunctional pain system or an ability to dissociate oneself from the pain of others.⁵ In brief, psychopaths have low fear because they have low pain. With regard to the somatic marker hypothesis, it is suggested that not only does the brain map and receive feedback from the body but that the same neural

⁵Intuitively, why should I fear that which cannot hurt me?

structures are involved in representing the self and the other—i.e. “shared neural representations” (Jackson, Meltzoff, & Decety, 2005, p. 777). The somatic-marker hypothesis in essence, claims that the brain is a captive audience of the body (it has no choice but to attend to the body). Insofar as bodily feedback is absent, the brain fails to make good decisions. MNS theory, insofar as it proposes that mere exposure to another’s actions or emotions activates bodily systems, implies that the primate brain is inextricably connected to its social group. Hence, if MNS is dysfunctional in psychopaths, social signals fail to evoke an appropriate response in psychopaths.

Viewing facial expressions triggers emotion even without conscious recognition (Adolphs, 2003). Keysers and Gazzola (2007) suggest that social cognition is made possible by two mechanisms: Theory of Mind (conscious reflection on the mental states of others) and simulation (a pre-reflective, bodily modeling of another’s affective state). If so, then the response perseveration and inability to extinguish a learned response as described by Newman et al. (1987) would make sense because psychopaths would be impaired in bodily simulation. Also, Blair’s Violence Inhibition Model would follow because the aversive signals given by a victim (such as when a psychopath is hitting another person) would not be accompanied by an experience of the victim’s emotion, even if cognitively, the psychopath knows what it means to be hurt.

Finally, how can MNS literature explain the language deficits of psychopaths? Rizzolatti and Arbib (1998) make the bold assertion that the human homolog of F5 in monkeys is Broca’s area—the area of the brain devoted to speech production. In their theory, language in humans has its origins in the primitive grammar of grasping and hand movements in monkeys. In their view, the F5 mirror neurons are capable of abstraction and can handle imperative and declarative linguistic structures which when combined with agents (e.g. self or other) is a pre-linguistic grammar of action.

This grammar of action is supposedly the precursor of verbal language which became dominant in humans. Seen in this light, the language deficits of psychopaths can again be linked to a failure of the social mirroring system whereby communication is facilitated by the involuntary transmission of affect and meaning via mirror neurons. Again, the implication is that linguistic meaning is not only transmitted via disembodied syntax (which psychopaths understand well enough) but also via bodily simulation.

In summary, the MNS literature sheds light on whether psychopathy is an adaptation or a disability. If it is shown that psychopaths have impaired MNS, should psychopaths have reduced culpability for antisocial actions? An analogy helps clarify this point: a person afflicted with leprosy does not feel pain and as a result, is prone to injure himself. We cannot blame the same person for being “careless” with himself if he sustains injuries to fingers or limbs. If psychopaths do not have the involuntary ability to feel the pain of another, we should not be as quick to apply the moral standard that is meant for healthy people. Contrary to expectation however, Fecteau et al.’s (2008) result suggest a more sinister view: psychopaths feel the pain of their victims but do the harmful action anyway. Once again, we raise the caveat that this finding needs to be replicated ⁶. That psychopaths have normal MNS activity bolsters the evolutionarily adaptive view of psychopathy which is discussed in the next chapter.

⁶I have gained the endorsement of a neuroscientist in Italy to do a similar project, provided that my postdoctoral fellowship grant application is successful.

Chapter 3

An Alternative View: Psychopathy as An Adaptation

Although the behaviour our evolved programs would, on average, have been adaptive (reproduction-promoting) in ancestral environments, there is no guarantee that it will be so now.

J. Tooby and L. Cosmides

3.1 Comment on the Sociobiology wars

Insofar as evolutionary psychology is influenced by sociobiology, it is important to justify an evolutionary approach such as that used in the present study because sociobiology is a fiercely debated topic. My evolutionary approach to psychopathy does not make claims about modules in the human brain/mind, but about the existence of psychopathy in the human population. In effect, the objective of the present investigation is much more limited than the goals of sociobiology, and independent of its problematic assumptions.

The core idea of sociobiology is that human social behavior evolves under Darwinian laws of natural selection (Buller, 2009; Wilson, 1975). In Darwin's view,

random mutations in organisms lead to trait differences conferring advantages. If there is a mechanism to transmit those differences to future generations, those mutations will spread in the population. Insofar as physical organs are concerned (i.e. the panda's thumb, bird wings, and the like), there is sufficient scientific evidence to infer the transmission of traits. Sociobiology extends the Darwinian approach to account for social behavior such as altruism—for example as to why vampire bats share blood (Wilkinson, 1984). According to Wilson (1975), the explanatory problem with altruism is that an individual reduces its own fitness in the act of sharing a resource with family or kin. But how can a reduction of fitness serve as a reproductive advantage?

Several explanations have been proposed such as inclusive fitness (Hamilton, 1964), and reciprocal altruism (Trivers, 1971). The former explains that a sibling is willing to share a scarce resource such as food because his “loss” is offset by increased reproductive success, since he shares half of his genes with the sibling. The latter explains why for example, in cleaner systems among fish, the host fish does not devour the little cleaner fish in its mouth. It is because for the host, the long-term benefit outweighs the marginal value of a meal (Trivers, 2002). On the part of the little fish, cleaning the host who visits a cleaning station is a stable supply of food. One major problem with sociobiology is that unlike physical organs where evidence comes in fossils and intermediate forms (e.g. Tiktaalik as a transitional form between fish and amphibian (Shubin, Daeschler, & Jenkins, 2006), there is a lack of physical evidence for a “selfish gene” as a means of transmitting selfish behavior. In particular, sociobiological claims about human nature are said to go beyond what is allowed by rigorous scientific evidence (Kitcher, 1985).

Like sociobiology, some practitioners of what is called “popular evolutionary psychology” are criticized for positing specialized brain modules (language for Pinker (2000), cheater-detection for Tooby/Cosmides (n.d.)) where there is a lack of neu-

roscientific evidence for the same. All that is available for direct examination are DNA strands, proteins, and brain tissue (Panksepp & Panksepp, 2000). In effect, the causal chain between physical mechanisms and the hypothesized function is absent. But while the existence of such specialized modules might be doubtful, it is accepted that the brain has “layers” with evolutionarily old emotion structures located in the caudal and medial areas of the brain (Panksepp, 2004; Damasio, 2000; MacLean, 1990). Therefore, the error made by some evolutionary psychologists is in their reification of brain functions, and not the overall evolutionary approach to the study human social behavior, as such. In fact, the same critics of “popular evolutionary psychology” hold that contrary to Tooby/Cosmides’ claim, socio-affective mechanisms did not arise in the Pleistocene era (i.e., Stone Age) but even earlier (Buller, 2009; Panksepp, 2004).

The present investigation, which is an empirical study of psychopathic behavior makes no specific claims on the following issues: (1) the presence of a psychopathic module in the brain; (2) the manner of transmission of psychopathic traits, be it genetic, epigenetic, behavioral or cultural. (See Jablonka and Lamb, (2005) for a comprehensive discussion of levels of selection.) There are plausible physical mechanisms underlying psychopathy such as a dysfunctional amygdala (Blair, 2007), defective mirror neurons (Fecteau et al., 2008), and a possibly differential expression of oxytocin and vasopressin (Panksepp & Panksepp, 2000), so my approach is not neuroscientifically unsound. But finding physical substrates for psychopathy and tracing their evolutionary origins is not my goal. Instead, my argument (see Chapters 6 and 7) is that since psychopathic traits are linked with greater defection in the Prisoner Dilemma, psychopaths will co-exist but not drive cooperators to extinction, because psychopathy is not an evolutionarily stable strategy. Conversely, psychopaths will not be driven to extinction. In sum, this investigation’s evolutionary approach does

not make statements about the brain, but about the state of psychopathy in the population.

3.2 Chapter Objective

In Chapter 2 psychopathy is explained from the lens of psychiatry with the mainstream view that it is a personality disorder rooted in dysfunctional brain systems. In this chapter, we review a different body of literature that uses the lens of evolution. This approach differs from the previous one on two counts. First, instead of an individual brain perspective, this approach takes the population perspective that is subject to evolutionary dynamics. These dynamics can be stated thus: more successful replicators (i.e. those that survive and reproduce better) tend to dominate a population compared to less successful ones in succeeding generations (Nowak, 2006). In time, less successful genes, ideas, or cultural forms go extinct. Secondly, the adaptive view takes a functional (as opposed to mechanistic) approach. Its major premise is that the brain evolved to solve problems related to living in groups (Tooby & Cosmides, 2009). Hence, studying psychopathy is at once a social and biological problem. The traits that we saw (e.g. manipulation, lying, lack of guilt) need to be seen in terms of what role they played in group living. Three mechanisms are required for this dynamic to happen: (1) there is variation in traits among replicators, (2) variation is tied to survival and reproduction; (3) a means of transmission of traits from one generation to another (Stearns & Koella, 2008). Seen in this light, psychopathy could be a more reproductively successful form that enables it to persist at a small fraction of the population. From a biological standpoint, fitness depends not only on reproductive success but also if those offspring are well adapted to their environment and therefore, survive long enough to reproduce.

In order for psychopathy as an evolutionary adaptive strategy to work, it needs

to be justified that psychopathy is both (a) subject to the dynamics of evolution, and (b) adaptive. We refer to these as the evolutionary and the adaptive parts of the argument. The evolutionary part is supported by evidence satisfying the three mechanistic requirements above. The adaptive part is supported indirectly by proxy measures, the most important of which are monetary winnings in laboratory games.

3.3 Psychopathy is subject to evolutionary dynamics

Enough evidence has been offered in chapter 2 about the first mechanism—psychopaths differ in brain systems, traits, and behaviors from non-psychopaths. It will now be argued that these expressions of psychopathy enhance reproductive success and are genetically transmissible.

3.3.1 Reproductive Success

Several lines of evidence support the argument that psychopaths are shrewd at mating. (1) *Psychopaths recognize the value of presenting a trustworthy image.* Male psychopaths present themselves differently to males and females. To the former, they present an image of promiscuousness and popularity while to females they present themselves as resourceful, trustworthy, and sincere (Tooke & Camire, 1991). The rationale is that females look for exactly those qualities in a mate. Knowing that those who look trustworthy get the most mating opportunities, psychopaths use a cognitive-behavioral maneuver to hide their promiscuous ways from women.

(2) *Psychopaths are “deceptive interpersonal strategists” that maximize short-term gain while minimizing long-term commitment* (Seto et al., 1997) *or a high ratio of mating effort (e.g. copulation, number of offspring) over parental effort* (Ellis, 1988). This hedonistic and irresponsible sexual strategy is indexed by age at first intercourse,

not using contraception in sex, and quick initiation of sex (vis-à-vis getting to know the mate well) from the time of meeting. Psychopaths are characterized by precocious and coercive sexuality and these traits are associated with antisocial behavior (Harris et al., 2007). In a study of sexual assault victims of psychopaths, Harris et al. observed that PCL scores correlated with the number of adult female victims but not with prepubescent victims. Importantly, those psychopaths who targeted children had lower PCL-R scores than those who did not target children. These findings suggest that psychopathy is driven by a genetic program geared towards reproductive success, and not just sexual pleasure for its own sake.

(3) *Psychopaths have the ability to spot the most vulnerable victims.* In a study where male students university students were asked to put themselves in the role of a mugger, those scoring higher in a self-report psychopathy scale (in particular, the Factor 1 dimension) were more accurate at judging the videotaped gait of other students for vulnerability (Wheeler et al., 2009). This research builds on previous findings that vulnerability to victimization manifests itself in manner of walking through indicators such as fluidity of movement, length of stride, and speed (Sakaguchi & Hasegawa, 2006; Murzynski & Degelman, 1996; Grayson & Stein, 1981).

In summary, the ability to assume a trustworthy façade, the ability to seduce partners soon after meeting them, and the ability to identify the most vulnerable victim are traits that confer reproductive advantage.

3.3.2 Genetic basis to psychopathy

The link between genetics and psychopathy is not as widely studied as the link between traits and behavior. Investigations carried out in this line have largely used twins to distinguish genetic from environmental effects. Taylor et al. (2003) assessed two cohorts of male monozygotic twins ($n = 270$) and dizygotic twins ($n = 128$)

for psychopathy using the Minnesota Temperament Inventory (MTI) ¹. Zygosity was determined by various means including physical similarity, biometric measures, and serological analysis. A strong correlation ($r = .74$) was found between genetic factors and scores in the MTI (Taylor et al., 2003). The same study found that about half of the covariation between the two psychopathy factors was genetic in nature (Taylor et al., 2003). A similar result was obtained by Blonigen et al. (2005) who found that the correlation between psychopathic traits was twice as large in monozygotic twins compared to dizygotic twins. This result is consistent with the view that psychopathy has a substantial genetic component and is heritable. A third study (Larsson, Andershed, & Lichtenstein, 2006) found that 63 percent of the variance in psychopathy scores was accounted for by genetic factors, with the remaining 37 percent of the variance accounted for by environmental factors (i.e. peer group). These three studies used three different measures of psychopathy but consistently implicated a genetic basis. That the samples used were community-based adolescent twins (instead of incarcerated samples) bypasses the theoretical difficulty of disentangling antisocial behavior and psychopathy.

3.4 Psychopathy is adaptive

3.4.1 Psychopaths Play a Rationally Competitive Strategy in the Prisoner Dilemma

As stated previously, the major drawback in laboratory games such as the Iowa Gambling Task and the Go-No Go paradigms is their asocial nature. Psychopaths, by definition, require social contexts for their manipulation, lying, and deceitfulness to be fully played out. The Prisoner Dilemma ² is a scenario in which two criminals,

¹The MTI is a unpublished 19-item scale authored by Loney et al. (Taylor et al., 2003) as part of the Minnesota Twin Family Study

²attributed to Albert Tucker who formalized the original ideas of two RAND Corporation em-

each one held in a different room, are interrogated by police. Each criminal ponders the following conundrum:

if I confess and the other stays silent, I go away free and he goes to federal prison. If he confesses and I remain silent, he goes free and I go to federal prison. If we both stay silent, we might get detained for a few hours but are soon freed for lack of evidence. If we both confess, we probably both serve a short sentence in the local jail.

Hence, the Prisoner Dilemma (PD), which explores cooperation and defection between parties is an excellent window into social decision-making by psychopaths. Unfortunately, only four studies are known to have been published using the PD with psychopathy or antisocial behavior.

The earliest known study involving psychopaths and the PD was done by Widom in the 1970s. In that experiment, primary and secondary psychopaths did not differ in level of cooperation or selfishness compared with normals (Widom, 1976). It should be noted that the psychopathy was estimated using the MMPI as the PCL had not yet been published yet at that time. In Spain, antisocial patients in a methadone maintenance program played a Prisoner Dilemma against a computer opponent. Compared with controls, antisocial individuals showed greater non-cooperation and retaliation when provoked by prior defection (Montañes et al., 2003).

A third study by Rilling et al. (2007) used fMRI to analyze the neural correlates of emotions (afraid, angry, guilty and the like) while playing a PD game. Brain activations were analyzed at the epoch when a choice was made and when the outcome was revealed. The two measures of psychopathy used were the Psychopathic Personality Inventory (PPI) (Lilienfeld & Widows, 2005) and the Levenson Primary and Ployees (Poundstone, 1993)

Secondary Psychopathy Scales (Levenson et al., 1995). Participants also rated their emotional reaction to the outcomes while in the scanner.

The important results of the study include: (1) Only in males (but not when males and females are combined) was there a significant negative correlation between psychopathy score and likelihood to continue cooperating (2) In aversive outcomes (i.e., those in which an individual cooperated while the partner defected), high psychopathic students showed lower activation in the amygdala compared to low psychopathic students, which indicate poor conditioning for aversive outcomes.

Recently, Mokros et al. (2008) had forensic patients who were committed to two high-security German hospitals played a PD experiment. Psychopathy was assessed using the PPI. Healthy subjects from the community were used as controls. Unlike the Rilling study, the PD game was presented in the context of a water rationing vignette in a drought-stricken village. The opponent was a computer programmed to play a tit-for-two-tats strategy, which means that the computer only plays defect after two successive choices of defect by the subject.

Mokros et al. (2008) concluded that: (1) Psychopaths played the non-cooperative choice more frequently than non-psychopaths and thereby earned significantly higher rewards. (2) Greater non-cooperation by psychopaths is linked with rebellious non-conformity, blame externalization, Machiavellian egocentricity and total score in the PPI. (3) Psychopathy is probably a beneficial strategy that is resistant to treatment because of higher rewards.

In summary, three out of four studies found psychopaths (or antisocial individuals) to be less cooperative in PD games. The most significant limitation in them is that only self-report measures of psychopathy were used. Although non-cooperation is inferior to mutual cooperation (and thus, is suboptimal), it is considered a rational choice in game theory because defection “immunizes” one from being exploited. In

terms of the evolutionary view, if one avoids the worst outcome (being exploited) and this action guarantees the same payoff as oneself to the opponent, then one prevents the other party from having greater fitness. In this sense, psychopathy is “defensively” adaptive.

3.4.2 Lying and Deception in Non-Psychopaths

If psychopathic traits are adaptive then healthy individuals (i.e. non-psychopaths) should sometimes exhibit psychopathic traits in their actions. I explore this line of thought by arguing that euphemistic language and common or conflicting interests in business are deceptive, dishonest, and exploitative.

Healthy individuals use self-deception to conceal facts that would otherwise cause subjective distress or cognitive dissonance. This kind of suppression manifests itself most clearly in language. In war, civilian targets are “collateral damage”, hostile combatants are “targets to be neutralized”, torture is “waterboarding” or an “enhanced interrogation technique”. “Self-deception lubricates the psychological machinery of slaughter, providing balm for an aching conscience.” (Smith, 2009, p. 109) Euphemistic language is part of a rationalization process that fades the moral aspect of a dilemma into oblivion (Tenbrunsel & Messick, 2004), hiding the conflict as if by magic, out of consciousness. Even when the outcome of an action (e.g. laying off workers) is the same, using terminology such as “cutting costs” hides the uncomfortable truth that people lose their livelihoods (or jobs, to use euphemistic language).

In the business world, the collusion of vested interests can lead to the deliberate falsification of information. There are many ways of making this happen all of which are based on the fact that accounting involves a lot of subjectivity such as in the timing for recognizing revenue, judging between a capital good versus an expense, and the like. Managers, whose performance is tied to quarterly results, have an incentive to

produce good numbers. Auditing companies are not entirely objective because they rely on other consulting business from the very companies they audit (Demski, 2003). Good accountants do bad audits not because they are morally corrupt but because of unconscious, self-serving biases (Bazerman et al., 2002). In short, where financial statements are supposed to be an accurate guide of a company's position, vested interests conspire to present a rosier-than-reality picture.

If common or shared interests can lead to collusion, are competing interests always a good thing? When there is a conflict of interest, what is rightful can take second place to money. Take the esteemed profession of medicine, for example. Physicians, are torn between providing the best treatment for patients based on scientific evidence and the lure of money from pharmaceutical companies. "Money" rarely comes in the form of outright bribes, but is packaged as "continuing medical education" which is none other than marketing in disguise (Relman, 2001). As a result, drawing the line between science and marketing becomes blurred. In an extreme case of callousness, deception, lying, and other psychopathic traits, Dr. Charles Nemeroff³, the former head of Psychiatry at Emory University was ousted for willfully concealing payments received from drug companies and promoting his own patent in a research publication (Armstrong, 2008). Evidence of other cases suggests that underdeclaring one's income (normally limited to US \$10,000) might be a widespread practice among researchers (Kaiser, 2009).

In summary, it is difficult to justify the claim that only psychopaths engage in lying, deception. Dilemmas exist in the business world that force individuals to choose between principles or money. Even if everyone took the moral path, cognitive

³In an unbelievable twist of events, the University of Miami hired Nemeroff to head its Psychiatry Department. The dean of the School of Medicine had this to say, "The Miller School of Medicine needs a strong leader in psychiatry and behavioral health and we have found that person in Charlie Nemeroff."

mechanisms are in place to suppress the ethical aspect of the dilemmas, reducing the choices to technical decisions. Even in psychiatry, the very field that gave rise to the concept of psychopathy, lying and deception take place. Whether the criterion is amount of damage, lack of compunction, or egoism, it is hard to argue that healthy individuals involved are not acting as if they were psychopaths themselves.

3.4.3 Lethal Aggression is Part of our Primate Heritage and Early Human Communities

In this subsection, the shared genetic legacy that humans have in common with primates and its implications for psychopathy is discussed. The logic of this approach is that the shared capabilities of extant species probably represent those of common ancestors (Byrne & Whiten, 1992). But since primates and humans still exist, this means that the traits we share with them (violence and deception) were likely selected for by nature. The literature in this controversial perspective is reviewed next.

Chimpanzees form male groupings and engage in territorial defense, raiding, and coalitionary killing (Wrangham & Peterson, 1996). Wrangham (1999) proposes in his chimpanzee violence hypothesis that adult males engage in violence only when there is a sufficient imbalance of power. Hence, there is an assessment involved—chimpanzees do not use brute force unconditionally, but only subject to a cost-benefit calculation. Chimpanzees have a repertoire of defensive and aggressive tactics, of which coalitionary killing is but one form. The evolutionary interpretation of lethal violence is simple: when the neighboring group is eliminated, the aggressors do not have to split food, territory, and females (Williams et al., 2004).

Reproductive success is the same motive identified behind early human communities. In studying the Yanomamö, Chagnon (1988) explains how violence is sustained by complex belief systems, rituals, and emotions. Anthropologists writing about

tribes often cite vengeance as a motive (Chagnon, 1988; Diamond, 2008). Vengeance contributes to the nexus of meaning that feeds a vicious cycle of retaliation in several ways. First, vengeance is connected with the experience of bereavement at the loss of a relative. The Yanomamö word is *hushuwo* which translates into “anger verging on violence.” (Chagnon, 1988, p. 986). Secondly, avenging the death of a relative confers on a man a reputation of *waiteri* or fierceness. Men who have killed are called *unokai*, signifying distinction. The ability to protect one’s tribe is a measure of a man’s value (Chagnon, 1988). Third, swift retribution deters future attacks and conversely, not avenging is a sign of cowardice and leads other tribes to steal their women.⁴ Finally, those who have killed have more wives and children than those who have not killed.

From an evolutionary psychology perspective, Buss and Duntley (2003) propose that aggression and homicide are evolved adaptations to the problem of acquiring resources, mates, and territory. Crucial to this argument is that aggression and homicide are not “blind instincts” but flexibly deployed, subject to numerical superiority, similar to the chimpanzee violence hypothesis of Wrangham.

In summary, violence and aggression serve the the adaptive purpose of enhancing survival and reproductive success. In chimpanzees, there is an assessment of costs vs. benefits of raiding a neighboring group. In humans, this cognitive aspect of aggression is reflected in vengeance which is still referred to colloquially as “getting even”.

3.5 How Game Theory Fits with the Adaptive View

The previous two subsections summarized research and theories arguing that psychopathy is not a deficit but is advantageous. The evidence is very interdisciplinary, ranging from genetic studies to the business world to anthropology. Game theory is a

⁴An example of placating the enemy is British PM Chamberlain’s signing the Munich accord with Hitler in WWII.

field that unites the evidence by providing a common analytical language. Whether the evidence comes from evolution which uses *reproductive success* or the business world or laboratory games which use *money* or its proxies, game theory treats them all as *expected utility*. From a game-theoretic point of view, the different kinds of evidence offered in support of the adaptive view of psychopathy are arguing that its traits provide greater expected utility compared to “normals.” The *evolutionary* aspect is simply Darwin’s theory that advantageous “mutations” yield better adapted phenotypes each succeeding generation.

An important question to ask is, how does the game-theoretic view mesh with the established deficit-centered theories? Deficit theories are mechanistic explanations while game theory is an abstract model—they operate at different levels. To use an example from Maynard-Smith (?), to explain why bird wings have the shape that they have, one has to take into account mechanistic variables such as drag, turning radius and the like. Analysis of this kinds simply looks at the individual bird, without taking into account the other members of the species. However, when the phenomenon of interest is the density of bird population, one has to allow for the behavior of conspecifics since variables such as mates and resource come into play. This latter phenomenon, is more appropriately handled in the language of game theory which weighs the costs and benefits of group living.

Returning to the phenomenon of interest in this project, psychopathy is typically studied from the individual brain approach (deficit theories) and not the population approach (adaptive view). This dissertation project takes the population approach and compares payoffs earned in games played for money by psychopaths and compares them to those of non-psychopaths. The questions and hypotheses of the present study are discussed in the next chapter.

Chapter 4

Methods

Strategic thinking is the art of outdoing an adversary, knowing that the adversary is trying to do the same to you.

A. Dixit and B. Nalebuff

4.1 Recruitment of Participants

The recruitment of participants involved two parts—screening and experiments. The screening part was done by administering the Levenson Self-Report Psychopathy (Levenson et al., 1995) (henceforth, LSRP) scale to a cohort of 1,412 first and second year Introductory Psychology students who signed up for course credit as part of the 2009 Carleton Mass Testing Program. Out of 1,412 test-takers, the scores of 8 students were excluded due to a high rate of “Decline to Answer” responses, yielding 1,404 valid test profiles made up of 956 females and 448 males. Since there is a known gender difference in psychopathy (Hare, 2003), only males were subsequently invited for the experiment portion. The entire participant recruitment and experimental protocols were approved by the Carleton University Ethics Committee (#09–140). The advert for the experiment in the SONA system appeared as shown in Table 4.1

The LSRP total scores were divided into quartiles to distinguish high scorers (top

Table 4.1: Advertisement in the SONA

Study name:	Personality Traits and Games Played for Money
--------------------	---

Abstract:	Are winnings in computer games related to traits?
Description:	Three games will be played where points won will be converted to real money. Experiment sessions will be played two individuals at a time. Aside from the chance to win additional money, participants will receive a guaranteed \$5 plus academic credit.
Eligibility:	Mass Testing Participation, Males only
Restrictions:	Must have completed Mass Testing 2009
Invitation code:	Participants must have a special password (invitation code) to sign up for this study. The researcher should provide this to them assuming they are qualified for the study.
Duration:	120 minutes
Percentage:	2 percentage
Researcher:	Sean Leach Email: sleach@connect.carleton.ca
Sign-up / Cancellation Deadline:	24 hours before the study is to occur
Study status:	Visible to participants (approved). Active study (appears on the list of approved studies)
IRB Approval #	09-140 (expires August 20, 2010)

quartile) from low scorers (bottom quartile). Of 448 male test-takers, 117 were in the top quartile and 110 in the lower quartile. These individuals were invited by e-mail to sign up for the experiments through the SONA system. The study was advertised as *Personality Traits in Computer Games Played for Money*. The experiments consisted of three games and one problem solving task administered on a computer terminal. The use of the SONA system for experiment sign-up minimizes the possibility of con- vivance between any two participants beforehand. A trained BSc Psychology Honours student served as experimenter and was blind to the profiles of the participants (i.e. whether they belonged to the top or bottom quartile). Participants were told that they would receive a guaranteed amount of five dollars plus an additional variable amount that depended upon their performance in the games. They were told that under no circumstance would have no out-of-pocket expenses even if they ended up with a net loss in the games. The experiments were run with two participants at a time.

4.2 Description of Instrument

The LSRP is a 26-item self-report instrument that is designed to measure psychopathic traits and behaviors in a non-institutionalized population (Levenson et al., 1995). Unlike the PCL-R that uses a combination of clinical judgment and behavioral records to establish a diagnosis, the LSRP takes an “endorsement” approach. Statements like, *Success is based on survival of the fittest* are presented to the test-taker who rates them on a four-point scale from *Disagree strongly* to *Agree strongly*. Using exploratory factor analysis, the authors reported two LSRP factors which correspond with the original two factor structure of the PCL-R. In developing the instrument, Levenson sought to correlate the LSRP with various measures of psychopathy and antisocial action, using Karpman’s distinction of primary and secondary psychopathy

as a guideline. Primary psychopaths are callous, manipulative, selfish and dishonest individuals with little emotion, while secondary psychopaths are anxious individuals who engage in antisocial actions under emotional distress (Karpman, 1948).

Lynam, Whiteside, and Jones (1999) did a confirmatory factor analysis which validated the two-factor structure of the LSRP. High scorers on the LSRP reported committing a greater variety of serious antisocial actions (e.g. taking a car, physical assault). Consistent with the findings of Newman regarding the passive avoidance deficit, high LSRP scorers are biased towards a reward-seeking and punishment insensitive response style. Brinkley et al. (2001) aimed to establish a correlation between the PCL-R and the LSRP scores using an inmate sample ($n = 549$). The correlations between the measures were low to moderate—0.35 for total scores, 0.30 for PCL Factor 1 and LSRP Factor 1, and 0.36 for the PCL Factor 2 and the LSRP Factor 2. The results likewise replicated the two-factor structure of the LSRP but the statistical fit was shown to be better for Caucasians than African-Americans. Like the Lynam study, Brinkley et al. (2001) also reported the association between psychopathy and passive avoidance.

In summary, the LSRP has been found to replicate the two-factor structure of the PCL-R and high scorers in it exhibit the same passive avoidance deficits observed in previous experiments. Given these features, the LSRP was deemed to be an appropriate instrument for measuring psychopathic characteristics in a sample of university students.

4.3 Physical Setting and Experiment Administration

All experiments were conducted in a quiet room equipped with three mice-equipped laptops connected to the internet. The connection to the internet was required by

two games which were web-based: the Prisoner Dilemma and the Tower of Hanoi. For the other two games, software was run locally on the laptops. The first laptop was used by the experiment administrator for recording experimental data.¹ The other two laptops were used by the participants in the three games and the control task. All responses in the games involved clicking the mouse to select from a set of choices. Participants were seated back-to-back as illustrated in Appendix B with no communication allowed all throughout the games. The games were run in a fixed order for all experimental runs in the following sequence: Prisoner Dilemma, Iowa Gambling Task, Cambridge Gamble Task, and Tower of Hanoi. The experiment administrator recorded the points won after each game (except for the Tower of Hanoi, in which moves and time were recorded) and at the end of the games, the points were converted to money. Participants were then debriefed, paid, thanked for their participation, and reminded not to reveal the details to other *Introductory Psychology* students.

4.4 Description of the Games

4.4.1 The Prisoner Dilemma

The first game was a web-based Prisoner Dilemma task that was played using *Econport* software (Cox & Swarthout, n.d.). In the Prisoner Dilemma, the payoffs depend on both one's choice and that of the other player (see Figures 4.1 and 4.2). The crux of the dilemma is that the best outcome for the pair is at odds with the best outcome for a single player. Referring to Figure 4.1, It is to both players' mutual benefit if they played {S, S}², yielding 2.5 points for each player. Observe however

¹i.e., Sean Leach.

²Instead of explicitly referring to the choices as "Cooperate" or "Defect", the names given are S, C, to guard against emotional bias due to the meanings. What counts in a Prisoner Dilemma game

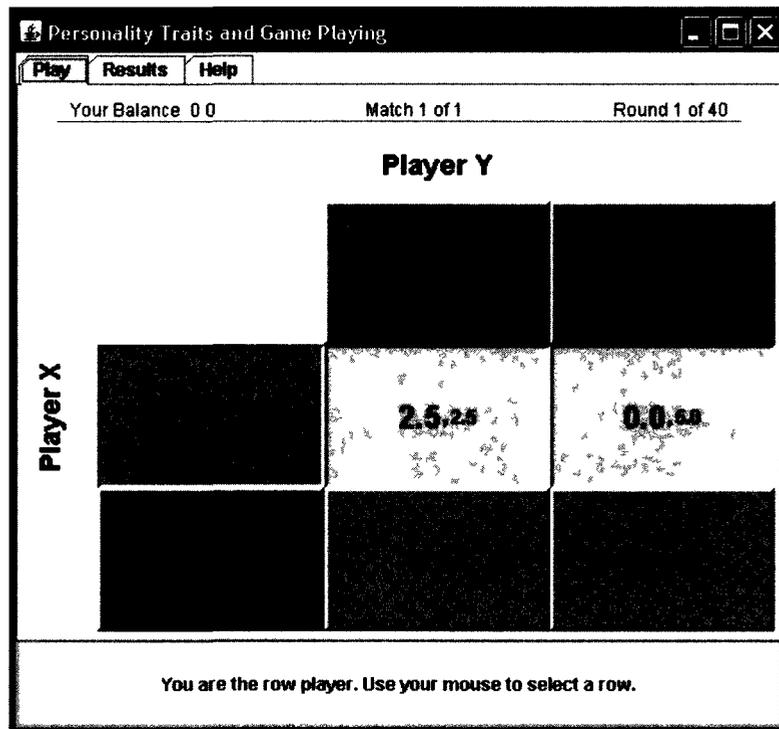


Figure 4.1: Prisoner Dilemma, Row Player plays "S"

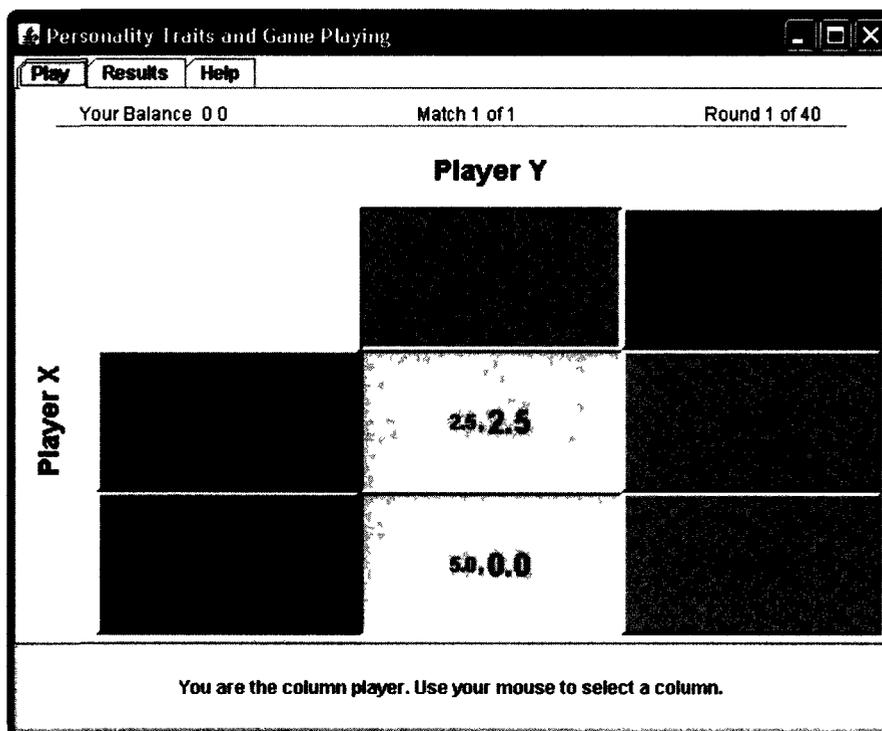


Figure 4.2: Prisoner Dilemma, Column Player plays “S”

that there is a temptation to defect to “C” which yields a superior payoff of 5 points, conditional upon the other player’s move of “S”. The Prisoner Dilemma was repeated for 40 rounds for each pair.

Participants were given the instructions for the game and to test for understanding, they were asked questions about the nature of the game. (See Appendix B on page 156) To familiarize them with the interface, they were asked to play 5 trial rounds. The real game consisted of 40 rounds. The points earned were accumulated and later converted to cash at the rate of 1 cent per 1 point won.

According to the adaptive hypothesis, it is more advantageous to be a psychopath among cooperators. We therefore have the following hypothesis in the Prisoner Dilemma.

Hypothesis 4.4.1: High psychopathic participants paired against low will earn more than high psychopathic paired against high psychopathic ones.

4.4.2 The Iowa Gambling Task

In the Iowa Gambling Task (IGT), participants are given a loan of 2,000 dollars and instructed to draw one card at a time from one of 4 card decks. The stated goal is to win as much money as possible. The experiment consists of 100 trials and there are 60 cards per deck. Two of the decks are “good” and lead to a net win in the long run while the other two are “bad” because they lead to a net loss. Consistent with the *response reversal* paradigm, the “bad decks” were initially conditioned to yield big wins which turn to big losses as time goes by. By contrast, the “good deck” payoffs got better as time went by although the wins were not as big. The primary measure of performance is the NET TOTAL (Bechara, 2007) defined as follows:

$$NET\ TOTAL = \frac{Cards\ from\ Good\ Decks - Cards\ from\ Bad\ Decks}{2}$$

are not the names of the strategies but the relative magnitudes of the payoffs.

As reviewed in sub-Section 2.4.3, individuals with psychopathic tendencies tend to have negative NET scores, i.e. they incur losses in the IGT and are insensitive to the change in punishment contingencies. We therefore have as our hypothesis:

Hypothesis 4.4.2: High psychopathic participants will have a lower NET score in the IGT compared to low psychopathic ones.

4.4.3 The Cambridge Gamble Task

In the Cambridge Gamble Task (CGT), a die is hidden in one of 10 red or blue cards and the player is to guess the color of the card that the die is under. He is to

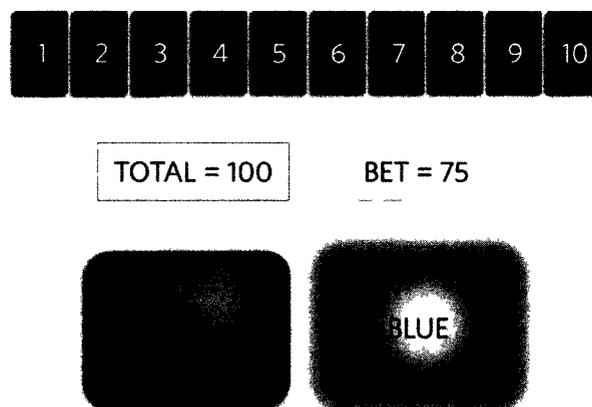


Figure 4.3: Cambridge Gambling Task

make this guess by pressing a red or blue button. (See Figure 4.3) After indicating his choice, the player assigns a bet that reflects his confidence in his choice. There are ascending and descending conditions in the order that the bets are presented: 5%, 25%, 50%, 75%, and 95% of his money. He starts with 100 points for each round and the win (or loss) is added to (or subtracted from) his earnings. The conversion rate is 1 dollar for every 1,000 points earned.

A recent study compared successful entrepreneurs in London's Silicon Fen clus-

ter of high-tech companies (Lawrence et al., 2008). Sixteen (16) entrepreneurs and seventeen (17) managers matched for age and intelligence participated in a Tower of London (also known as Tower of Hanoi) task and the CGT. Entrepreneurs had founded at least two companies while managers were non-founders. The Tower of London involves “cold” decision-making-one that does not involve judgments of risk and reward. The CGT on the other hand, involves “hot” decision-making, because the participant stands to gain or lose points depending on their bet. The results showed no group difference in “cold” decision-making but a significant difference in winnings in favor of the entrepreneurs as a result of greater impulsivity and risk-taking. We therefore have the following hypothesis:

Hypothesis 4.4.3: High psychopathic participants will earn more money in the CGT compared to low psychopathic ones.

4.4.4 The Tower of Hanoi

The Tower of Hanoi Task was administered in order to control for executive function. In this game, there are three towers. The left-most tower has a certain number of rings. The rings are to be moved one at a time from the left tower to the right tower. The goal is to use the smallest number of moves to accomplish the task. The exact instructions given for the task were:

- Move all the rings to the right-most tower. Click to take a ring, click again to put a ring.
- You can only take a ring at a time.
- A ring cannot be put on a smaller ring.

Two levels of difficulty were used: 4 rings and 5 rings. See Figure 4.4 for a screenshot of the 4-Ring version. A trial round with 3 rings was used to train the

participants in the task. With the Tower of Hanoi serving as a control task, no specific hypothesis was tested but the number of moves and the time taken to complete the task were recorded.. The software used to run the task is available at: <http://lbalbuen.solidwebhost.com/hanoi.html>.

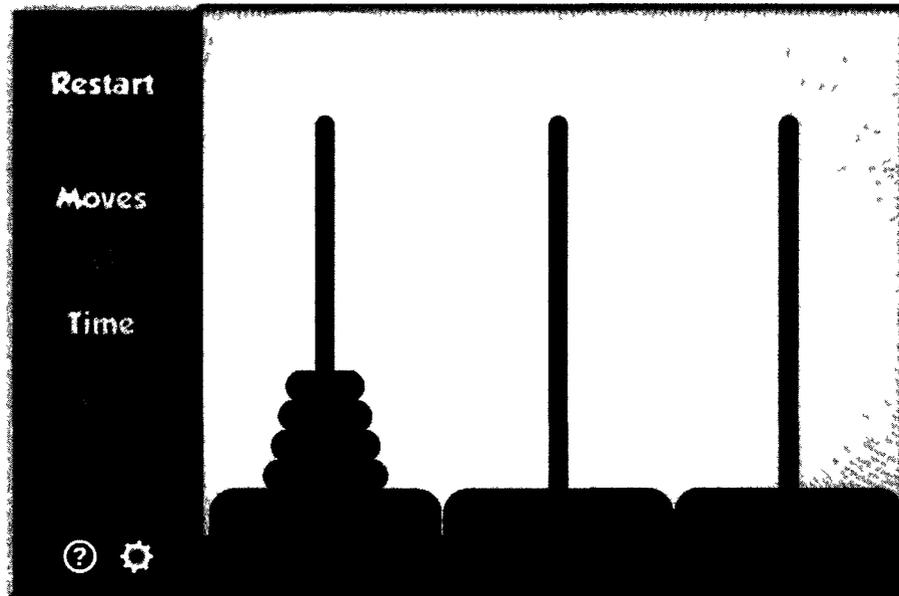


Figure 4.4: The Tower of Hanoi task, 4-Rings

4.4.5 Debriefing

After each game, the monetary winnings of the participants were tallied by the experimenter in a spreadsheet. After the three experiments and the Tower of Hanoi were completed, the participants were debriefed and told about the true purpose of the study. In accordance with guidelines of the American Psychological Association, the participants were given the choice to keep their data private. None of the participants withdrew their consent after the debriefing. The accumulated winnings in the games were then converted to money at the rates given in Table 4.2.

Table 4.2: Cash Conversion Table

	Virtual Points	Cash equivalent
Prisoner Dilemma	1	1 cent
Iowa Gambling Task	1,000	1 dollar
Cambridge Gamble Task	1,000	1 dollar
Tower of Hanoi	N/A	N/A

Chapter 5

Results

5.1 Participant Characteristics and Overall Winnings

A total of 48 subjects (26 High P, 22 Low P) took part in the three experimental games. There were no significant differences in age or Tower of Hanoi scores. Psychopathy groups scored significantly differently in LSRP scores. Table 5.1 provides mean scores across groups for LSRP, age¹, and the Tower of Hanoi task.

Table 5.1: Mean LSRP scores, Age, and 4-Ring Tower of Hanoi Scores (SDs) as a Function of Low vs. High Psychopathy Groups

	Low P (n = 22)		High P (n = 26)			
	Mean	S.D	Mean	S.D	t	p
LSRP	45.68	1.99	69.27	4.39	23.26	< .001
Age	21.14	6.14	18.96	1.18	-1.64	0.12
4-Ring Hanoi Moves	30.05	14.6	29.42	16.04	-0.14	0.89
4-Ring Hanoi Time	68.14	45.86	63.42	48.53		0.73

Note: LSRP = Levenson Self-Report Psychopathy

¹Although age did not vary significantly, it was suggested by Dr. Brook that it would be interesting to study the time course of psychopathy scores if they decline over time. The present study is cross-sectional in design so the hypothesis could not be tested.

Because the experiment was advertised as games played for money, the points won (or lost) in each game are important performance measures in themselves. The monetary winnings themselves are not as important because the payouts were scaled so as to be affordable. In none of the games was there a significant difference between High P and Low P. CGT (Descending treatment) was the closest to achieving significance where High P earned more than Low P. Descriptive statistics are presented in Table 5.2.

Table 5.2: Descriptive Statistics of Points Won Across Groups for the Prisoner Dilemma, Iowa Gambling Task and Cambridge Gamble Task

	Low P		High P		t	p
	Mean	S.D	Mean	S.D		
Prisoner Dilemma	76.13	23.52	74.83	18.56	0.23	0.82
Iowa Gambling Task	-427.95	1183.88	-244.38	1308.27	0.50	0.62
Cambridge Gamble Task (Asc)	1985.73	762.1	1948.31	743.2	0.17	0.86
Cambridge Gamble Task (Desc)	1017.75	216.98	2576.65	505.32	1.8	0.08

Note: Asc = Ascending, Desc = Descending

In the Prisoner Dilemma, $n = 42$; Iowa Gambling Task, $n = 44$;

Cambridge Gamble Task, $n = 48$

5.2 Prisoner Dilemma

A total of 42 students ² played the Prisoner Dilemma under two (2) treatment conditions: *Mixed Pairs* and *Homogeneous Pairs*. In the former, subjects are paired against a different *Subject Type* and the same type in the latter. In all, there were seven pairs each of High–High, Low–Low, and High–Low based on Levenson Scores.

The primary variables of interest were *Rate of Cooperation* (i.e., count of “Coop-

²Not all 48 subjects participated in the Prisoner Dilemma because in some experimental runs, only one subject showed up.

erate” moves) and *Winnings*.

5.2.1 Descriptive Statistics and Time Trend

Rates of Cooperation stratified by *Subject Type* were both approximately normally distributed, Shapiro–Wilk’s $W_{HighP} = 0.93$, $p = .14$ and Shapiro–Wilk’s $W_{LowP} = 0.92$, $p = .11$. The distribution of cooperate moves is displayed graphically in Appendix E. As expected, *Rate of Cooperation* did not correlate with *Winnings*, Pearson’s $r = .18$, $p = .25$. This is due to the structure of the game where cooperating pays more only if the other party also cooperates. Overall rates of cooperation and winnings in the 40 rounds are summarized in Table 5.3, stratified by *Subject Type* and *Match Type*.

Table 5.3: Descriptive Statistics by Subject Type \times Match Type for Cooperation and Winnings in the Prisoner Dilemma

	No. of ‘Cooperate’ moves		Winnings	
	Low P	High P	Low P	High P
Subject Type				
Match Type	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Homogeneous	14.43 (5.97)	11.57 (7.14)	76.64 (24.14)	66.57 (14.10)
Mixed	21.86 (9.41)	18.71 (12.67)	75.64 (24.08)	91.36 (15.49)

Figure 5.1 shows the overall decline of cooperation over time by match number. Of the 21 matches, 16 showed a negative slope with time, 4 had a flat slope and interestingly, 1 match had an increasing slope with time. In the Figure 5.1 plots, there are separate regression lines for each participant in each match color coded according to *Subject Type*.

5.2.2 Predictors of Rate of Cooperation and Points Earned

A univariate General Linear Model (ANOVA) was conducted to determine if *Rate*

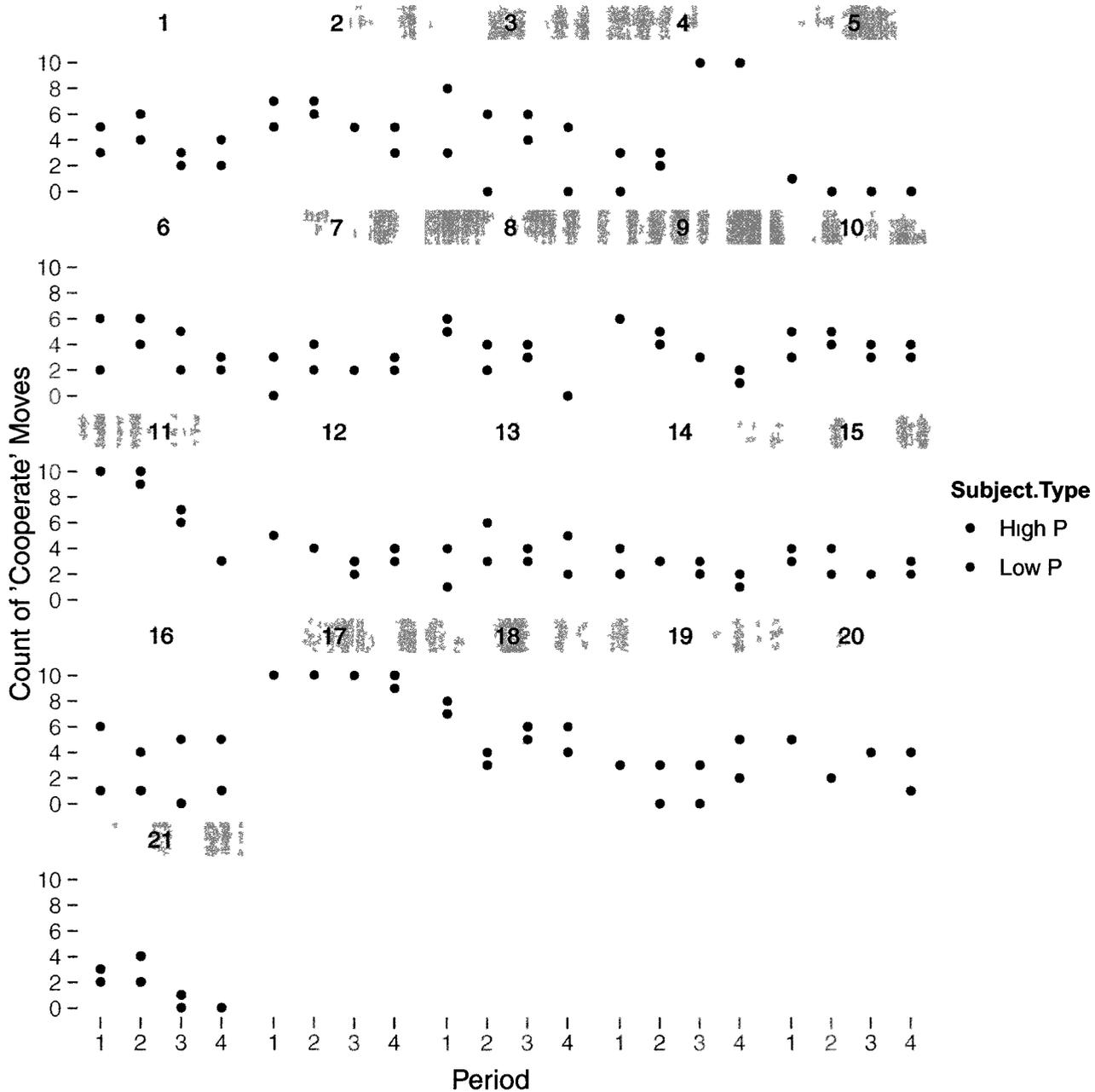


Figure 5.1: Cooperation Time Trend by Match

of *Cooperation* could be predicted by a linear combination of the following predictor variables: *Subject Type*, *Match Type*, and *Subject Type* \times *Match Type*. A separate ANOVA was performed with *Winnings* as the dependent variable and the same predictor variables³. With *Rate of Cooperation* as DV, the main effect of *Match Type* was significant, $F(1, 38) = 7.18, MSE = 68.98, p = .011$. Neither the main effect of *Subject Type* $F(1, 38) = 1.22, MSE = 68.98, p = .28$ nor the interaction term *Subject Type* \times *Match Type* $F(1, 38) = .003, MSE = 68.98, p = .96$ were significant. With *Winnings* as DV, the interaction term *Subject Type* \times *Match Type* was very nearly significant, $F(1, 38) = 3.91, MSE = 396.87, p = .055$ while none of the main effects were significant. For *Subject Type*, $F(1, 38) = 0.19, MSE = 396.87, p = .67$ and for *Match Type*, $F(1, 38) = 3.33, MSE = 396.87, p = .076$. The results of these two ANOVAs can be visualized in Figures 5.2 and 5.3.

To summarize, the ANOVA had the following results: (1) *Rate of Cooperation* differed by *Match Type* but not by *Subject Type*. Both High P and Low P subjects were more cooperative when paired against each other than when paired with their own kind. The claim is not that mixed pairing “caused” greater cooperation since none of the subjects participated in more than one treatment. Still there is a treatment effect that warrants further investigation. We shall return to this matter in the Chapter 6. (2) While Low P participants earned almost the same amount regardless of *Match Type*, High P participants earned significantly more in the mixed condition. This finding went exactly as predicted in Hypothesis 4.4.1. High P subjects cooperated less when paired with each other, hence pairing with a more cooperative partner led to the expected result.

There are two limitations of the analysis conducted in this subsection. First, the

³a MANOVA was initially considered so as to protect against Type I error, but a significant Box’s M result showed that the equality of covariance assumption was violated

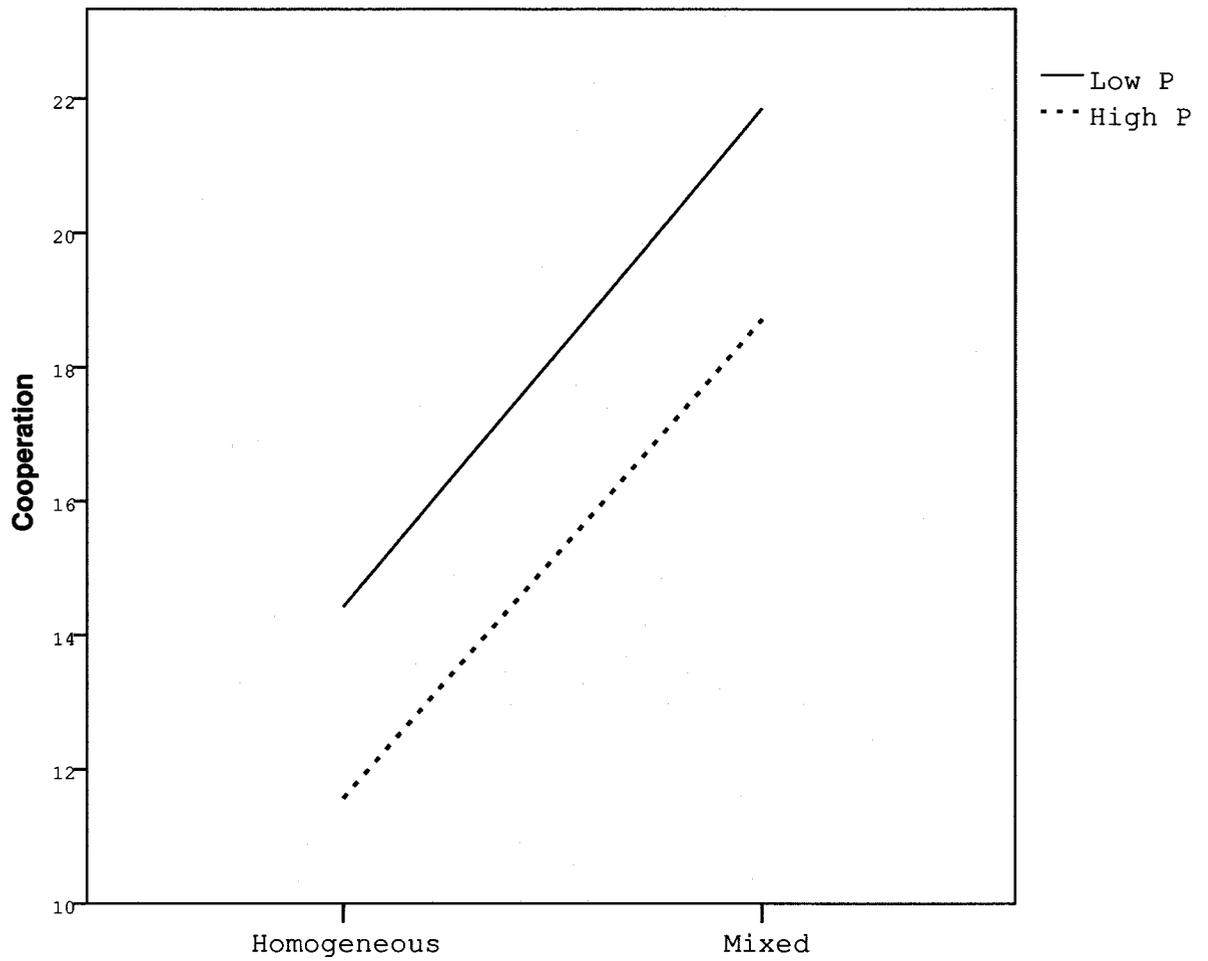


Figure 5.2: Mean Rate of Cooperation in the PD by Match Type

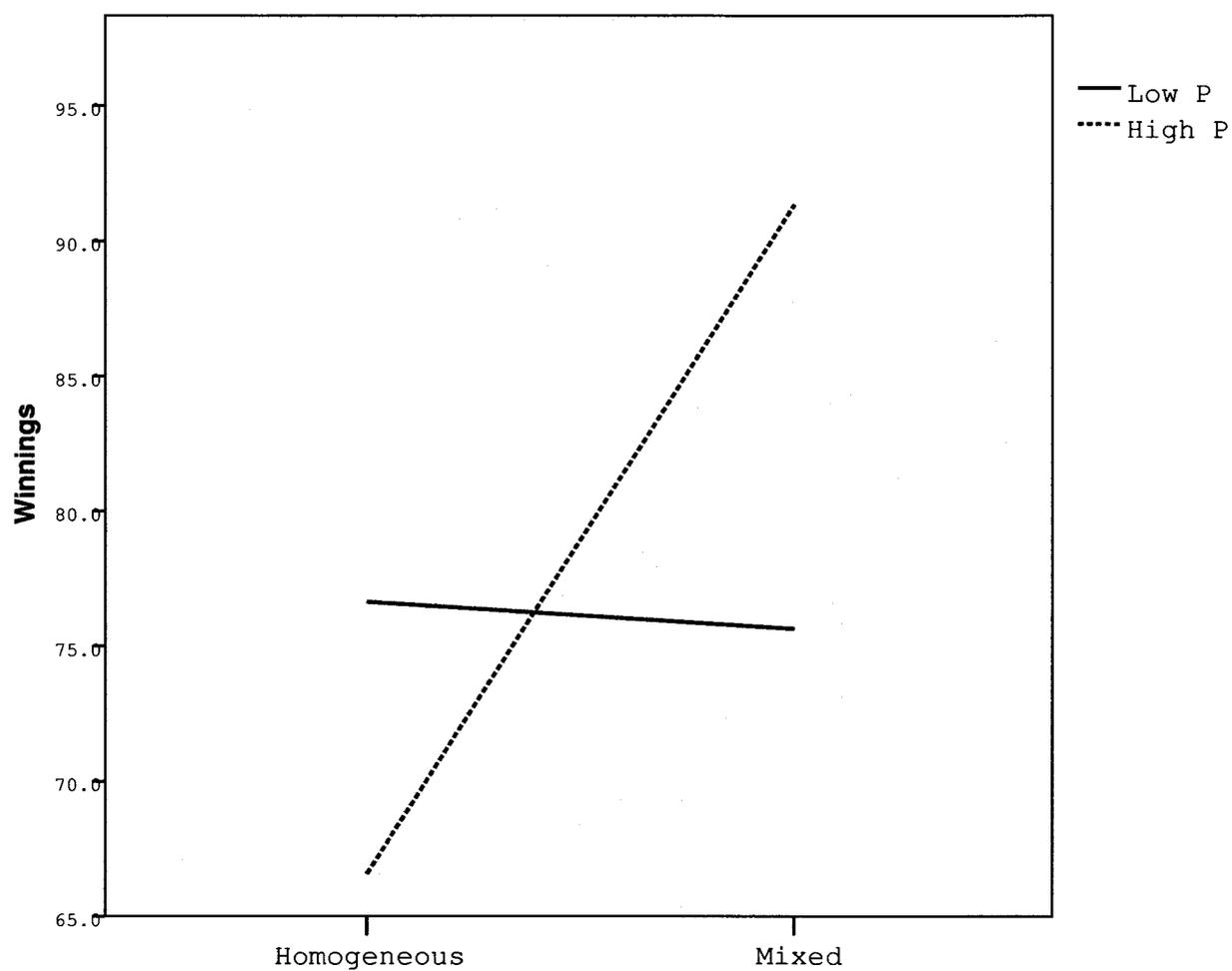


Figure 5.3: Mean Winnings in the PD by Subject and Match Type

DVs (e.g. *Cooperation Rate* and *Winnings*) were measured without considering the move of the other player in the pair. Remember that there were 21 matches of two players each, but that each player only played against a single opponent for the entire 40 rounds. But by its very nature, the Prisoner Dilemma is a social game where one adjusts one's choices according to the previous choice of the other player. Secondly, it would be good to simulate how each player's strategy would fare against every other strategy. In the next subsection, we address the first limitation by considering cooperation and defection in the context of the choices of the other player in the pair. In Chapter 6, we attempt to address the second limitation by modeling a generalized contest where one strategy plays against every other strategy.

5.2.3 Greedy, Generous, and Reciprocal Strategies

In the 1980's, Axelrod hosted a Prisoner Dilemma tournament of computer programs (Axelrod & Hamilton, 1981). Contestants submitted strategies for 200 trials and each strategy played every other strategy. In the first tournament, 14 strategies were submitted, some of which used a Markov process or Bayesian inference. The winner of the tournament was Tit-for-Tat (TFT), submitted by Anatol Rapoport. TFT always cooperates in the first round and copies the competitor's previous move every round thereafter. Other possible strategies in the finitely repeated PD are "Always Defect", "Always Cooperate", "Alternate C and D" and so on. A second round of competition was announced and this time 62 entries from six different countries were submitted. Once again, TFT by Anatol Rapoport won the tournament.

Because of the large number of possible strategies in the repeated PD, it is useful to classify strategies into different kinds. One such typology has three categories: **Greedy**, **Generous**, and **Reciprocal** (Johansson, Carlsson, & Boman, n.d.). A **Generous** strategy is one in where the count of (C,D) exceeds that of (D,C), where

the notation (C,D) means “I cooperated and the other player defected” in a given round. Similarly, (D,C) means “I defected and the other person cooperated” in a given round. A **Greedy** strategy is just the opposite—the count of (D,C) exceeds that of (C,D), which means that I had more (D,C) outcomes than (C,D) outcomes. A **Reciprocal** strategy, as its name suggests, has about the same count of (C,D) as (D,C) ones.⁴ Observe that different from the previous section, now the choice of the other player is taken into account, with the limitation that sequential dependency (i.e. the effect of a previous move on a succeeding move) is still not considered.

Johansson’s categorization was applied to the results of the present investigation. The distribution of strategies is summarized in Table 5.4. Note that aside from *Match Type*, the other two variables are ordinal in nature. *Subject Type* is determined by scores in the Levenson pretest while *Strategy Type* (**Generous**, **Reciprocal**, **Greedy**) coding increases with the number of DC rounds. Fisher’s exact tests were conducted to test for association.⁵ For the Homogenous matches, there was no link between *Subject Type* and *Strategy Type*, $p = .81$, Fisher’s exact test, 2-sided. The opposite was true in the Mixed matches where there was a link between the two variables, $p = .045$, Fisher’s exact test, 2-sided.

Table F.1 in Appendix F tallies the count of each outcome per player. From this table, we can calculate a measure, referred to as *Self-advantageous Defection Index*, which is simply the ratio of DC to CD outcomes for a given player⁶. It is reasonable to think that the *Winnings* over 40 rounds would increase with *Self-advantageous*

⁴(C,C) and (D,D) outcomes are not counted in this analysis as the intent of the analysis is to measure the asymmetry between how much I “screwed” the other player as opposed to how much I got “screwed”.

⁵Because of small cell sizes, Fisher’s exact test is used instead of chi-square. Thanks to Dr Katharine Kelly for this point.

⁶For four players, with a zero count in either CD or DC, .1 was added in the cell frequency, so as not to exclude them from the analysis. For the *Self-advantageous Defection Index* graph, two of the logged values are extreme outliers and are not shown in the graph for aesthetic reasons

Defection Index. This is indeed the case and there is a strong correlation between the two variables, Pearson's $r = .83, p < .01$. Figure 5.4 shows the linear relationship of

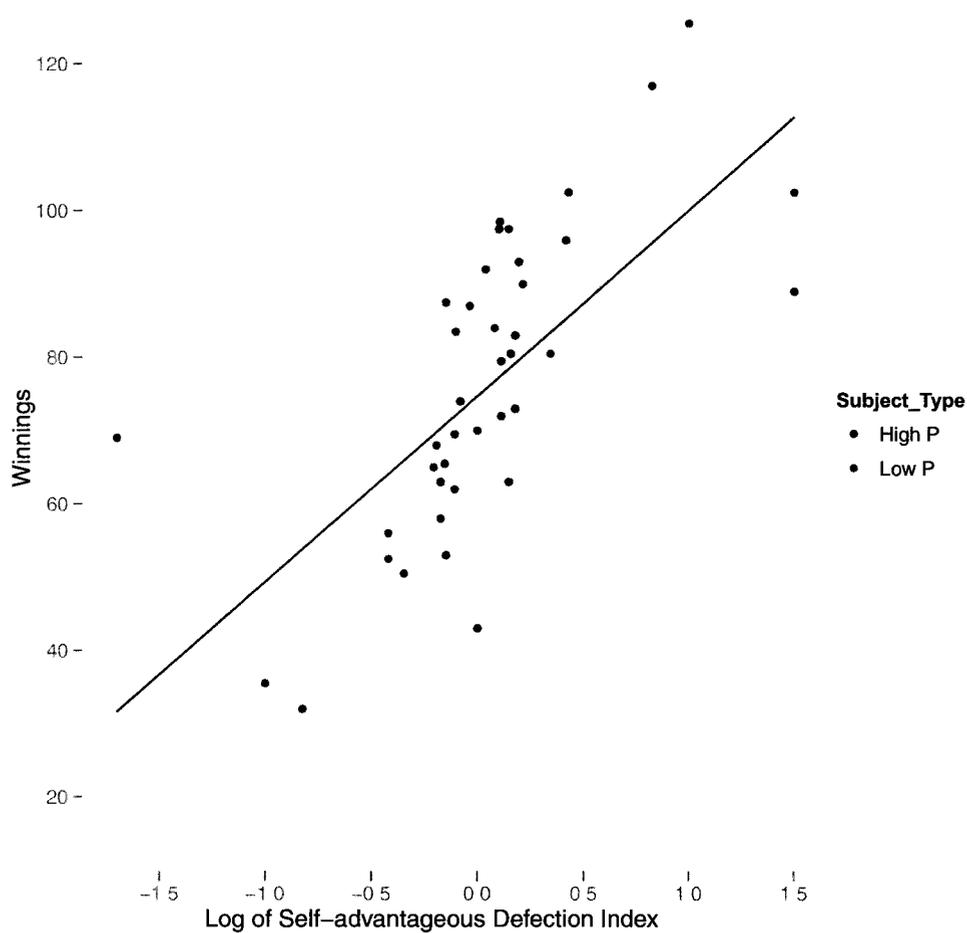


Figure 5.4: Scatterplot of $\text{Log}(\text{Self-advantageous Defection Index})$ and *Winnings*

the log-transformed *Self-advantageous Defection Index* with *Winnings*. Interestingly, there is no clear separation between the *Subjects Types* in the scatterplot (red vs. blue dots).

Table 5.4: Distribution of Strategies

Match Type		Strategy Type		
		Generous	Reciprocal	Greedy
Homogeneous	Subject Type			
	Low P	5	4	5
	High P	4	6	4
Mixed	Subject Type			
	Low P	3	4	0
	High P	0	4	3

5.2.4 Section Summary and Comparison to Previous Research

The ANOVA revealed that *Rate of Cooperation* was significantly different between *Match Types*. With *Winnings* as DV, the *Match Type* \times *Subject Type* was very nearly significant ($p = .055$). While Low P subjects won the same amount regardless of *Match Type*, High P participants won significantly more in the *Mixed* matches. This jump in winnings was enabled by a propensity for *Greedy* strategies in High P when playing against Low P. The result of the present investigation is consistent with three previous studies (Mokros et al., 2008; Rilling et al., 2007; Montañes et al., 2003) that found a higher defection rate among more psychopathic subjects. The present study is unique for two reasons: (1) real monetary rewards was at stake (vis-à-vis, water rationing for a drought stricken village in Mokros (2008); and (2) subjects played against a known human opponent (vis-à-vis computers in Montanes (2003) and Mokros (2008)). While previous studies focused on the link between psychopathic traits or neural correlates and PD playing styles, the present study also includes an analysis of the rationality of selfish playing styles in Chapter 6.

5.3 Iowa Gambling Task

Of the original 48 subjects, two participants from each *Subject Type* were excluded from the analysis because they picked from the good decks for the entire first half (50 trials)⁷ Of the remaining 44, 24 subjects belonged to the High P group while 20 belonged to the Low P group. Performance in the IGT did not differ in terms of proportion of good moves, NET scores, and reaction times, save for reaction times in one period. These are summarized in Table 5.5. In this section, the High P and Low P patterns of results on three measures are modeled: (1) Net Total Scores, (2) Moving Average of Good Deck picks, and (3) Streakiness for Good and Bad choices. The results are then compared to previous literature on the IGT.

5.3.1 NET Scores

The NET Total Score for the entire 100 trials did not vary by *Subject Type* with a mean score of 16.58 ($SD = 28.18$) for High P subjects, as compared with a mean score of 12.50 ($SD = 27.83$) for Low P subjects, $t(42) = .48$, $p = .63$. As mentioned in Subsection 4.4.2, the NET Total Score indicates how many more cards were drawn from good decks than from bad decks, with negative scores indicating more cards drawn from bad decks. NET Totals were also computed for five periods of 20 moves each using repeated measures ANOVA with *Time* as the within-subjects variable and *Subject Type* as the between-subjects variable. The interaction of $Time \times Subject Type$ was not significant, $F(2.94, 123.35) = 0.12$, $MSE = 92.89$, $p = .95$.⁸ The time course of Period NET Total scores is shown in Figure 5.5.

Two remarks have to be made regarding the main effect of *Time*. First, note

⁷It is very likely that they had prior knowledge about the IGT whether from reading or leaked information from earlier runs of the experiment.

⁸Mauchly's test indicated that the sphericity requirement for treatment levels was unmet $\chi^2(9) = 32.68$, $p < .001$, hence, Greenhouse-Geisser-corrected degrees of freedom are reported here.

Table 5.5: Descriptive Statistics in the IGT Across Groups

	High P (n = 24)		Low P (n = 20)		t	p
	Mean	SD	Mean	SD		
Bad Deck Picks						
Trials 1 to 20	11.96	3.29	12.40	4.77	0.36	0.72
Trials 21 to 40	8.96	3.71	9.75	3.89	0.69	0.49
Trials 41 to 60	8.21	4.4	8.85	5.00	0.45	0.65
Trials 61 to 80	6.21	4.99	6.00	4.95	0.14	0.89
Trials 81 to 100	6.38	5.59	6.75	5.38	0.22	0.82
Net Score						
Trials 1 to 20	-3.92	6.58	-4.80	9.55	0.36	0.72
Trials 21 to 40	2.00	7.46	0.50	7.78	0.65	0.52
Trials 41 to 60	3.67	8.76	2.30	10.00	0.48	0.63
Trials 61 to 80	7.33	10.14	8.00	9.90	0.22	0.83
Trials 81 to 100	7.50	11.04	6.50	10.76	0.30	0.77
Reaction time (ms)						
Trials 1 to 20	848.25	369.99	872.16	424.19	0.42	0.68
Trials 21 to 40	569.23	294.24	584.00	296.39	0.37	0.72
Trials 41 to 60	494.13	320.60	588.61	407.92	2.18	0.03
Trials 61 to 80	397.50	187.01	425.85	271.48	0.96	0.34
Trials 81 to 100	450.17	327.08	430.22	230.94	0.40	0.69
% of Trials from Good Decks	Count	%	Count	%	χ^2	p
Trials 1 to 20	233	49%	192	48%	0.03	0.86
Trials 21 to 40	305	64%	245	61%	0.49	0.45
Trials 41 to 60	323	67%	263	66%	0.23	0.63
Trials 61 to 80	351	73%	304	76%	0.95	0.33
Trials 81 to 100	362	75%	288	72%	1.32	0.25

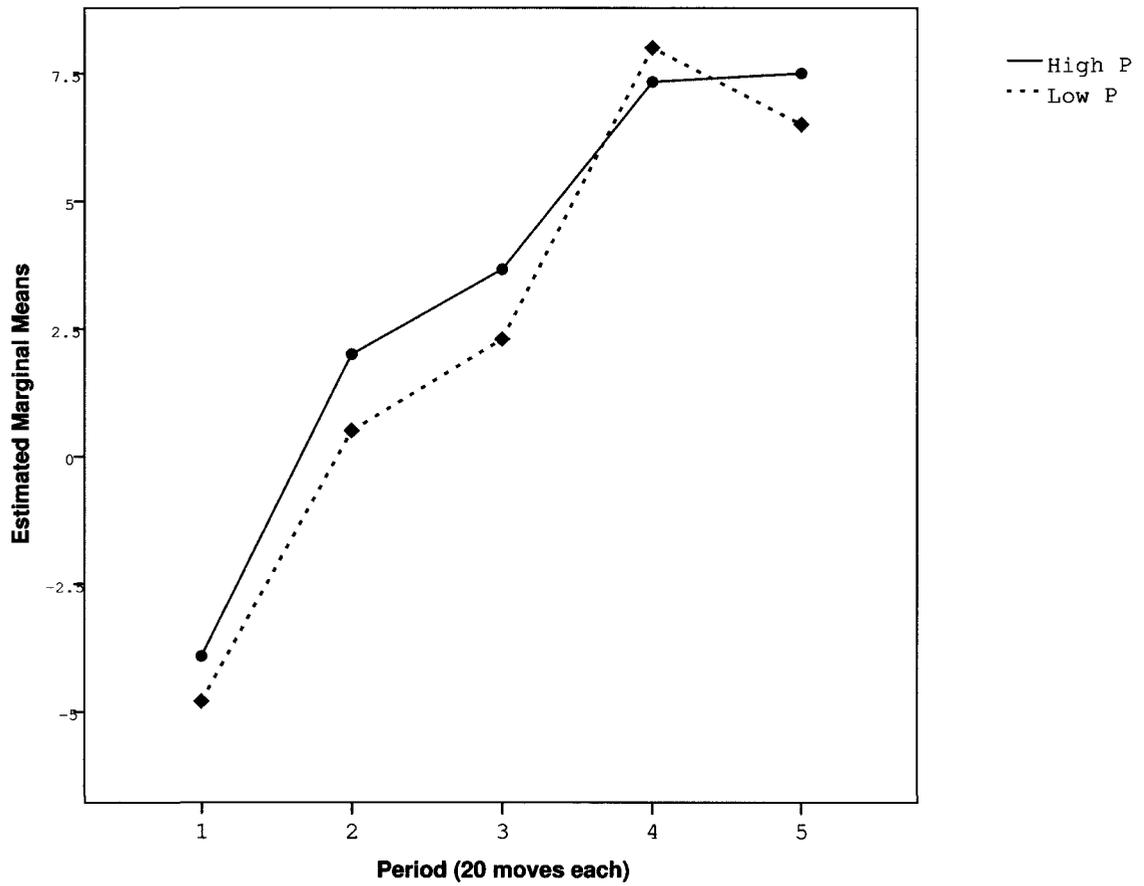


Figure 5.5: Net Total Scores in the IGT Between Groups by Period

that in the first period, the Net Scores for both groups are below zero (i.e., most of the cards drawn were from bad decks) but they climb past zero starting in period 2, implying that subjects learn to favor the good decks by the second period (moves 21–40). Secondly, both the linear and quadratic components of *Time* were significant. For the linear trend of *Time*, $F(1, 42) = 31.21$, $MSE = 118.67$, $p < .001$ and for the quadratic trend, $F(1, 42) = 4.59$, $MSE = 62.55$, $p = .038$. Neither the cubic or quartic components were significant.

The significant quadratic *Time* trend implies that the NET Score means do not increase monotonically, but plateau or decrease at some point. There is a plateau effect for High P subjects at Period 5 and a decrease at Period 5 for Low P subjects. Repeated measures ANOVA calculates NET Score at discrete time points, in this case, as the mean for a 20–move period. To determine whether the frequency of good choices varied by *Subject Type* at any point during the 40 trials, a moving average plot was created in the R language (R Development Core Team, 2009).

5.3.2 Moving Average of Good Deck Choices

The R package `freqMAP` (McCulloch, 2010) was used to construct a moving average of the frequency of good deck choices, stratified by *Subject Type*. A moving window of 10 moves was used. In addition to plotting the mean, `freqMAP` also calculates a 95 percent confidence interval around the mean. Unlike in the repeated measures analysis, the dependent variables of interest were: (i) the number of cards drawn from good decks (instead of NET Scores), and (ii) the frequency of winning cards drawn. The moving average of good choices is in Figure 5.6 and the moving average of wins is in Figure 5.7.

As shown in Figure 5.6, the frequency of good choices did not vary by *Subject Type* at any point and neither did the frequency of wins as shown in Figure 5.7.

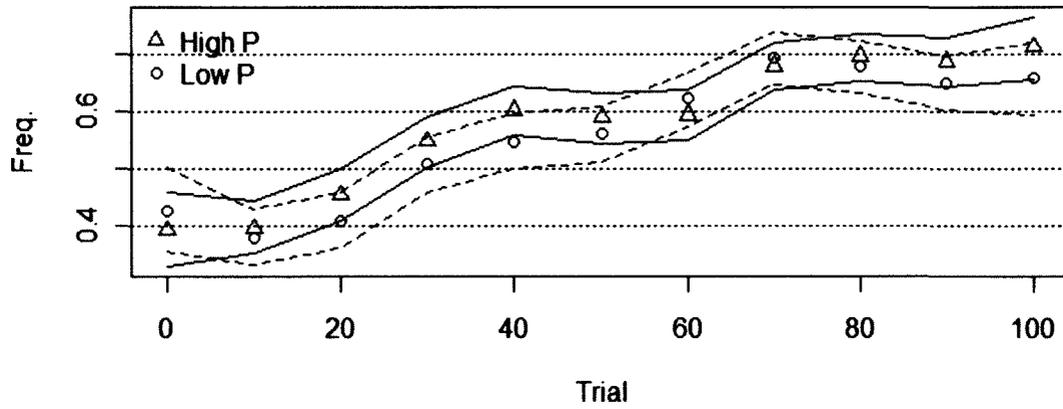


Figure 5.6: Moving Average of Good Deck Choices

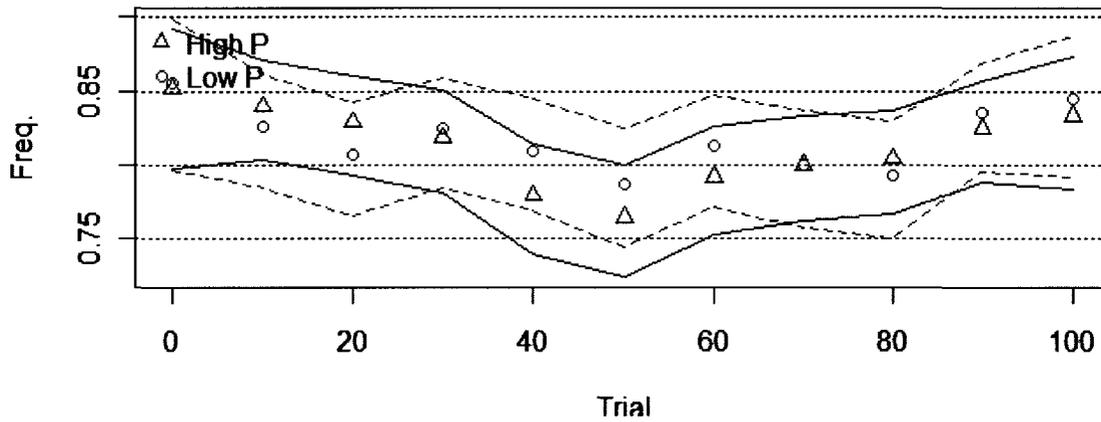


Figure 5.7: Moving Average of Wins

Player B's choice pattern exemplifies what is called the "independent, identically distributed" model where the relative frequency of choices is constant all throughout the 100 trials. Stated differently, the probability of picking from a good deck is independent of the previous pick. Player A by contrast is "streaky" because he goes through stretches of G's and of X's. While both players have the same frequency of good picks, Player A's longest streak of G's is 11 while Player B doesn't have a streak at all.

For each of the participants by *Subject Type*, each string of choices was examined for streaks of G's (Good Streaks) and X's (Bad Streaks). See Appendix G for the streaks of choices. A computer program in VB.Net was written to count the number of streaks per *Subject Type* for each half of the 100 trials. Counting was stratified by half to determine whether participants became more or less streaky by half. The VB.Net program produced streak lengths for its output. These raw streak lengths were then binned in 6 category lengths: 2 to 3, 4 to 5, 6 to 7, 8 to 9, 10 to 11, and 12 and above. Figure 5.8 presents the result graphically.

Streakiness can serve as a measure of learning in the following way. Once it is known that good decks are linked with net wins, then G choices should be streaky—call this *perseverance*. Alternatively, if bad decks are known to be disadvantageous but there are still long streaks of X choices, then there is *perseveration*. Hence, the learning over time question can be formulated thusly: do participants become streakier in good choices and less streaky in bad choices in the second half of the IGT trials as compared to the first half? The between-groups comparison can also be recast thusly: are High P participants more perseverative and less persevering than Low P participants?

Before presenting the answers to these questions, the logic behind the analytical test for streakiness which is a permutation test is sketched. Each participant has a

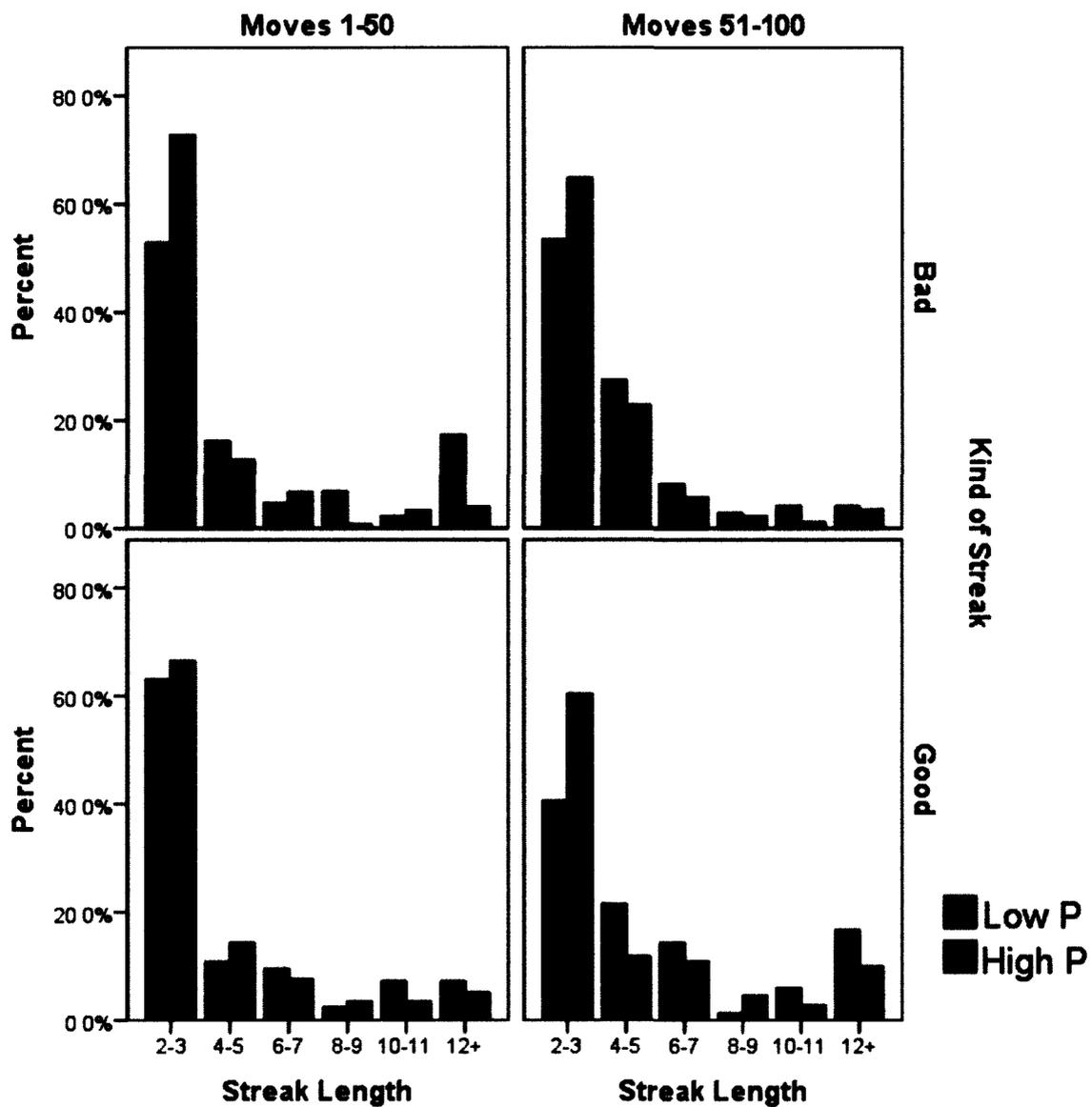


Figure 5.8: Streakiness in the IGT

string of choices like XXXXGGGGG or GXGXGXGG with a length of 50 for each half. The streak lengths counted by the VB.Net program are first segregated by *Subject Type*. The mean length of streaks for High P was then calculated—call this μ_H . Then the streak lengths for both groups were combined and 1,000 possible subsets were drawn from the combined set and their mean length recorded—call this μ_S . μ_H and μ_S are then compared for statistical significance. This technique falls under an area of statistics called resampling and was implemented using the R package `twot.permutation` (Maindonald & Braun, 2009). The main reason for using the `twot.permutation` is that the distribution of streaks departs substantially from the normal (as shown in Figure 5.8) distribution and that a t-test of mean streak length on the original data would therefore be inappropriate.

5.3.3.1 Bad Deck Streakiness

In this subsection, we focus only on the streaks of X's which index perseveration. We call this variable *bad deck streakiness* and compare it between participant types. The `twot.permutation` with 1,000 simulations revealed a statistically significant difference in bad deck streak lengths between subject types in the first 50 trials, $p < .01$ but not in the second half, $p = .36$. Contrary to expectation, Low P participants had longer bad streak lengths overall. The descriptive statistics for bad deck streakiness are summarized in Table 5.7.

To examine possible reasons for the surprising result with bad deck streakiness, the proportion of losses coming from bad decks was analyzed. Recall from Table 5.5 that there was no difference in bad deck or good deck picks between *Subject Types* in any period and this was confirmed by the moving average plot. Could it be that the proportion of losses coming from bad decks was different between subjects? Surprisingly enough, the High P participants had higher proportion of losses coming

from bad decks in both first and second halves compared to Low P. The descriptive statistics for bad deck losses are summarized in Table 5.8.

Table 5.7: Descriptive Statistics of Bad Deck Streak Lengths Across Groups based on 1,000 simulated Permutations, Mean (SD)

	First Half		Second Half	
	Trials 1 - 50		Trials 51 to 100	
	Mean	SD	Mean	SD
Low P	5.68	5.16	4.19	3.12
High P	3.72	3.26	3.73	3.47
t	10.15		0.99	
p	< .01		0.36	

Table 5.8: Descriptive Statistics of Bad Deck Losses Across Groups by Half, based on 1,000 simulated Permutations, Mean (SD)

	First Half		Second Half	
	Trials 1-50		Trials 51-100	
	Mean	SD	Mean	SD
Low P	5.85	2.56	4.75	3.48
High P	9.13	1.26	9.46	2.27
t	3.28		4.71	
p	< .001		< .001	

Because High P subjects were “disfavored” by the IGT software, the count of losses coming from bad decks was entered as a covariate ⁹ in a repeated measures ANOVA with *Mean Streak Length* as dependent variable, *Subject Type* as the between subjects factor and *Half*(1 vs. 2) as the within subjects factor. The main effect of *Half* was significant, $F(1, 40) = 11.37, MSE = 20.21, p = .002$. The main effect of

⁹The principle behind adding covariates in a linear model is to equalize subjects on a variable that could possibly confound the dependent variable measure.

Subject Type was significant $F(1, 40) = 7.40, MSE = 6.14, p = .01$. The repeated measures ANOVA confirms the `twot.permutation` test result. The change in bad deck streakiness with time can be seen from the plot in Figure 5.9

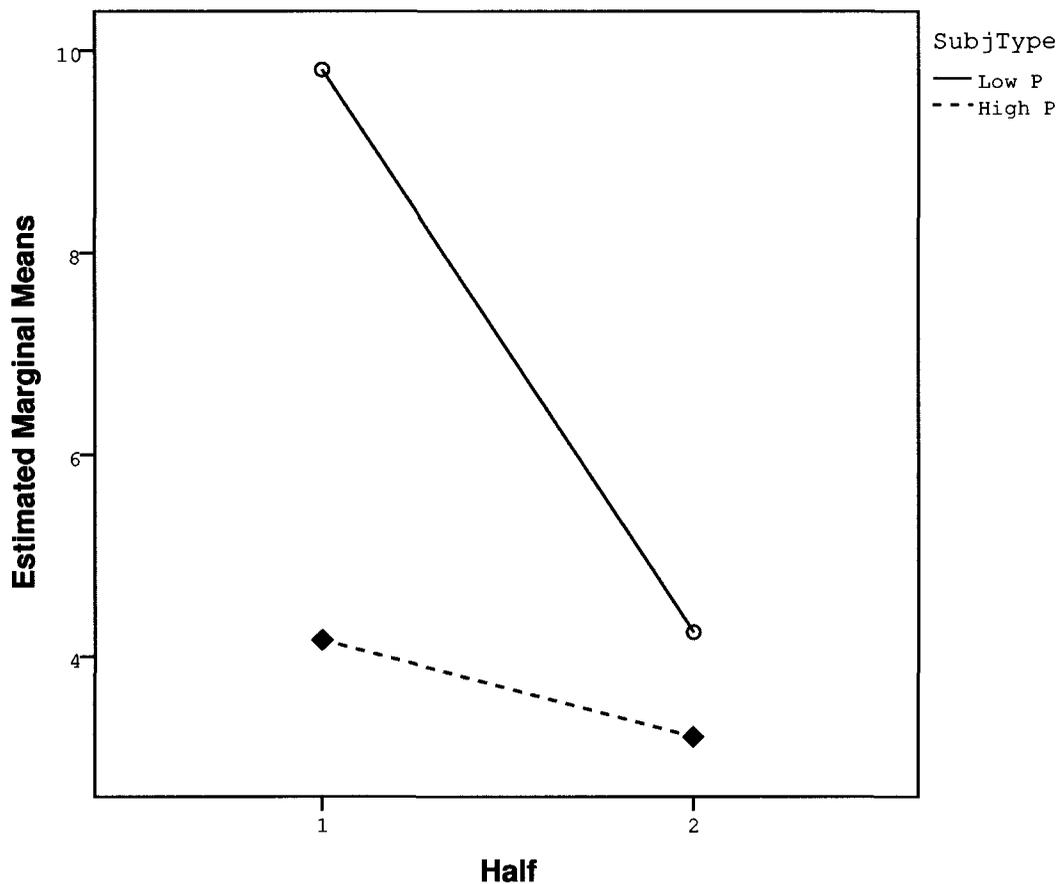


Figure 5.9: Bad Deck Streakiness, with Covariates 1^{st} Half Losses (Centered) = $-.0036$, and 2^{nd} Half Losses (Centered) = $-.0218$

5.3.3.2 Good Deck Streakiness

In this subsection, we focus only on the streaks of G's which index perseverance. We call these runs good deck streakiness and compare this variable between partici-

part types.

High P participants did not differ from Low P participants in good deck streakiness in the first half, $p = .33$ and in the second half, $p = .45$. High P participants were significantly more streaky in good choices in the second half compared to the first half, $p = .03$. Similarly, Low P participants were also more streaky in the second half compared to the first, $p = .03$.

The descriptive statistics for good deck streakiness are summarized in Table 5.9.

Table 5.9: Mean Length (SD) of Good Streaks Across Groups by Half based on 1,000 simulated permutations

High P		Low P	
Moves 1-50	Moves 51-100	Moves 1-50	Moves 51-100
Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
4.33 (4.82)*	6.48 (9.35)*	5.10 (5.88)[§]	7.43 (7.95)[§]
* Mean difference was significant, $p = .03$			
[§] Mean difference was significant, $p = .03$			

To summarize this subsection, the concept of streakiness reveals that although the frequency of bad choices did not vary by *Subject Type*, in the first 50 trials, Low P subjects were significantly more perseverative. From first to second half, perseveration by Low P subjects decreased significantly. By contrast, High P subjects were already non-perseverative in the first half and this did not change in the second half. This can be explained by higher losses sustained by High P in bad decks. In terms of good choices, High P and Low P participants became significantly more persevering in the second half.

5.3.4 Section Summary and Comparison to Previous Research

In this section, evidence supporting Hypothesis 4.4.2 that High P subjects would

get lower NET Scores than Low P subjects was not found. On the contrary, we found that both subject types learned to favor the good decks starting in Period 2. The moving average of good choices confirmed the ability of both subject types to reverse initial learning. In the streakiness analysis, Low P participants were more streaky for bad choices in the first half but were less so in the second half. High P participants were less streaky for bad choices in the first half and this did not change in the second half. Both subject types became more consistent with good choices in the second half of the IGT.

The lack of association between psychopathy and risk-taking (i.e. bad deck preference) in this study is consistent with previous IGT results by Schmitt and Newman (1999) but different from three other studies (Honk et al., 2002; Mitchell et al., 2002; Blair et al., 2001). The present study differed from the previous ones in two ways: (1) We used the Levenson Self-Report Psychopathy instrument, and (2) our study sample was made up of college students, instead of male offenders or incarcerated individuals. In this way, it might be more appropriate to compare our result with IGT results conducted in a normal population. One such study (Glicksohn, Naor-Ziv, & Leshem, 2007) reported that even in a non-clinical setting, 46 percent of individuals exhibit poor performance as measured by a lack of increase in preference for good decks over time. In the current study, only 32 percent ($\frac{14}{44}$) showed a flat or decreasing linear trend over time. There was no association between *Subject Type* and *Learning* (dichotomized as 'Yes' or 'No', depending on the slopes of lines in Figure 5.10), $\chi^2(1) = .79, p = .38$. Nine High P subjects¹⁰ and five Low P subjects¹¹ had a flat or decreasing linear trend. In Figure 5.10, subject numbers ending with an 'H' are

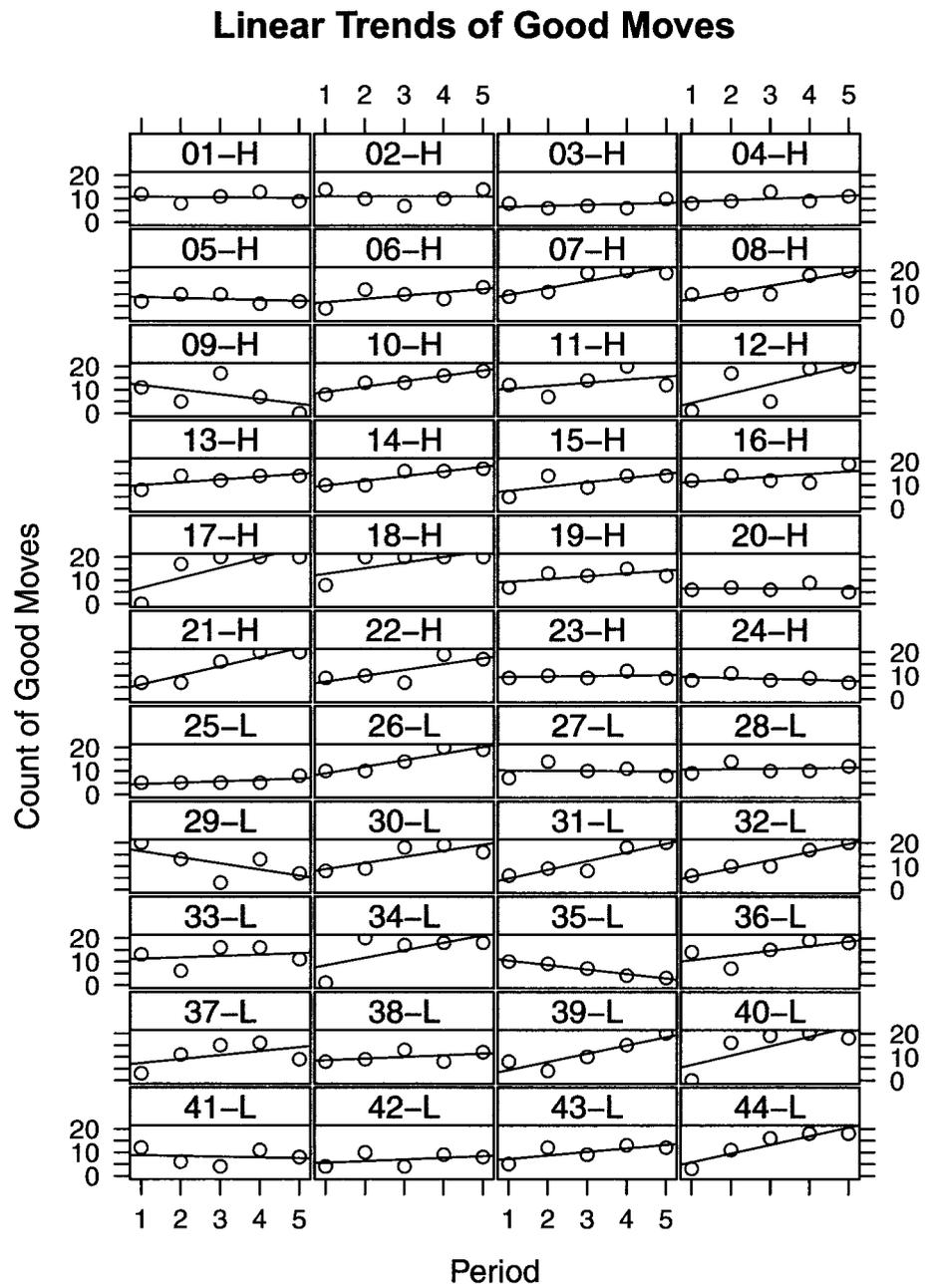


Figure 5.10: Individual Subject Trends of Good Choices in the IGT

High P while those ending with an ‘L’ are Low P. Using Bechara’s (1998) criterion that “normal performance” is indexed by the choice of greater than 50 cards from the good decks, 30 out of 44 participants met or exceeded this threshold. The learning ability in the IGT of our sample compares favorably against those of previous studies, especially those with incarcerated samples.

5.4 Cambridge Gamble Task

As discussed in the literature review, the CGT is designed to examine the role of impulsiveness and risk preference in decision-making. In this section, the results were analyzed on two levels: the overall sample analysis which suggested patterns of scores, and a restricted sample analysis consisting of the highest Levenson scorers (HHP)¹² and the lowest scorers (LLP). The rationale for this method is that differences that were found marginally significant or weak in the overall sample were amplified in the subsample.

5.4.1 Entire Sample Analysis

In the overall sample ($n = 48$), psychopathy was treated as a continuous variable. Thus, Levenson scores were tested for association against impulsivity and risk preference variables. The raw scores of several dependent variables departed substantially from normal so non-parametric techniques were used.

Impulsivity. The rank-ordered Levenson scores correlated weakly with both the log transformed *choice latency* time (Spearman’s $\rho = .08, p < .01$) and log transformed *response latency* time (Spearman’s $\rho = .06, p < .01$). *Choice latency* is the interval between the deck presentation and the subject’s choice of a color. *Response*

¹⁰1-H, 2-H, 3-H, 4-H, 5-H, 9-H, 20-H, 23-H and 24-H

¹¹27-L, 28-L, 29-L, 35-L and 41-L

¹²as a mnemonic, HHP stands for “Higher High P”, i.e., the upper half of the high quartile, while LLP stands for “Lower Low P”, i.e., the lower half of the low quartile

latency is the interval between the choice of a color and the time that a bet is placed. Thus, the finding of *longer* latency among the *more* psychopathic subjects went against theory and expectations.

Risk Preference. Interestingly, rank-ordered Levenson scores also correlated with the average amount bet for all trials (Spearman's $\rho = .06, p < .01$). To test preference for risk when the chances of winning and losing were equal, the analysis was restricted to those trials where the ratio of red to blue was 1:1 (i.e. there were 5 red and 5 blue boxes). Levenson scores were associated with higher bets, Spearman's $\rho = .13, p = .01$. However, when the psychopathy was treated as a dichotomy, High P subjects placed marginally significantly higher bets compared to Low P subjects, *Wilcoxon's* $W = 15,987, p = .054$. This finding motivated the restricted analysis using the subset of extreme scorers.

In summary, analysis at the overall sample level suggested that even though High P subjects took higher risks, they also spent more time to presumably assess those risks as reflected by longer latency times. Further analyses were conducted with extreme scorers as a subsample.

5.4.2 Restricted Sample Analysis

5.4.2.1 CGT Summary Measures

The procedure for selecting the sub-sample involved taking the highest and lowest quartiles of the sample of 48 individuals using their Levenson scores. Recall that these 48 individuals were themselves highest and lowest quartiles of 1,412 Carleton *Introductory Psychology* students who participated in the 2009 Mass Testing Program. Twenty-seven (27) subjects, made up of 11 HHP and 16 LLP composed the subsample in subsequent analyses. The mean Levenson score of HHP was 73.27 ($SD = 3.67$)

and for LLP it was 44.69 ($SD = 1.08$). To put these scores in perspective, a score of 73 ranked in the top 3 percent of the male SONA test-takers, while a score of 45 would be in the 92nd percentile.¹³ The mean age for LLP was 21.50 ($SD = 6.95$) and for HHP it was 18.82 ($SD = 1.17$) which were not statistically different, $p = .15$. Subject characteristics and CGT measures are summarized in Table 5.10.

Table 5.10: Subject Characteristics and CGT Measures

	LLP		HHP		Comparison	
	Mean	SD	Mean	SD	<i>t</i>	<i>p</i>
Subject variables	n = 16		n = 11			
Age	21.5	6.95	18.82	1.17	1.51	0.15
4-Ring Hanoi Time	76.38	51.13	60.55	58.52	0.75	0.46
4-Ring Hanoi Moves	33.00	16.03	27.27	17.49	0.88	0.38
4-Ring Hanoi Moves per min	0.53	0.24	0.54	0.18	0.22	0.82
Levenson Score	44.69	1.08	73.27	3.67	25.14	< .001
Impulsivity-Related						
Choice latency (ms)	1829.22	1525.8	2299.13	2076.59	5.53	< .001
Response latency (ms)	990.06	1062.24	1062.11	1017.55	1.4	0.15
Delay aversion	0.17	0.12	0.27	0.12	2.08	0.048
Deliberation time	404.67	101.17	462.67	139.5	2.77	0.010
Risk Preference						
Overall proportion bet (%)	0.52	0.12	0.55	0.06	0.74	0.46
Quality of decision-making	0.95	0.15	0.91	0.08	0.73	0.47
Risktaking	0.57	0.12	0.61	0.08	0.96	0.35
Risk adjustment	2.44	1.04	1.64	0.71	2.21	0.037
Total Points won						
Ascending Condition	2071.31	763.83	2016.36	698.57	0.11	0.91
Descending Condition	2460.16	1120.18	3988.00	3095.39	1.83	0.08

*Note: LLP = lower half of Low Psychopathic students;
HHP = higher half of High Psychopathic students;*

¹³These figures are for male scorers only. About two-thirds of the SONA test-takers were women.

The CGT includes several measures each for impulsivity and risk preference. For the former, the scores of interest are: (1) *choice latency* and (2) *response latency* as described previously; (3) *delay aversion*¹⁴, which is the degree to which a subject has difficulty in withholding a response; and (4) *deliberation time*, which is the mean latency time for those trials where the boxes were not equal in number. Two-sided t-tests revealed that the subgroups differed in *delay aversion*, $t(25) = 2.08, p = .048$, in *deliberation time*, $t(25) = 2.77, p = .01$, in *choice latency*, $t(1,798) = 5.53, p < .01$, but not in *response latency*, $p = .15$. Thus the restricted analysis confirmed the longer *choice latency* time of more psychopathic subjects found in the overall sample analysis. Paradoxically, more psychopathic subjects took longer to make choices but were also more averse to delays.¹⁵

In terms of *Points Won*, HHP and LLP earned roughly the same amount in the Ascending trials. In the Descending trials, HHP's higher earnings were marginally significant ($p = .08$). Importantly, participant winnings were converted to cash and participants were aware of this throughout the games. This experimental feature made the gambling more closely approximate real life behavior.

With regard to risk preference, CGT provides the following scores of interest: (1) *risk-taking*, which indicates the percentage of points gambled when the boxes are unequal and the subject chose the majority color; (2) *overall proportion bet*, which unlike *risk-taking* includes all assessment trials; (3) *risk adjustment* which measures how the pattern of betting changes as the asymmetry of box ratios increases; and (4) *quality of decision-making*, which is a count of the proportion of all trials in which the subject chose the majority color. The subgroups did not differ in *quality of decision-*

¹⁴Calculated as the mean proportion of the current points that the subject chose to risk on gamble test trials for which they had chosen the more likely outcome.—from CANTAB unpublished documentation.

¹⁵This can be accounted for by the fact that delay aversion contrasts *Ascending* and *Descending* modes while choice latency is an overall measure.

making, $p = .47$, *risk-taking*, $p = .35$, *overall proportion bet*, $p = .46$, but differed in *risk adjustment*, $t(25) = 2.21, p = .04$. The difference in *risk adjustment* indicated that LLP were better at aligning their bets with the odds of winning.

5.4.2.2 Modeling Subgroup Betting Patterns

The main dependent variable in the CGT is the percentage of money staked (*Stake*) as the ratio of colored boxes (*BoxRatio*) changes. A plot of the *Stake* with *BoxRatio* suggested an interaction effect of *Subject Type* and *Order of Stake Presentation*. In particular, while the percentages staked in the Ascending order did not differ, more psychopathic subjects tended to place higher bets in the Descending condition when the chance of winning was low. In Figure 5.11, note the non-overlapping confidence intervals in the Descending condition when the box ratio is roughly equal (i.e., 5:5 or 6:4).

To further examine the nature of the dependent variable, a linear mixed model (Pinheiro & Bates, 2000) analysis was fit to the data. Linear regression models are of the form

$$Y_i = \beta_0 + \beta_1 X + \varepsilon_i \quad (5.1)$$

where the intercept β_0 and slope β_1 are fixed effects and X (or more probably, a vector of X's) are predictor variables. The subscript i is an index for the subject. The equation contains a single error term ε_i for each subject, which when aggregated becomes σ . A linear mixed model (LMM) differs from linear regression models in two ways: (1) The intercept and slope in an LMM are themselves outcome variables (Singer & Willett, 2003), which are regressed against second-level Z variables. For this reason, an LMM is hierarchical, with the top-level equation called Level 1, and the lower-level equations called Level 2. (2) The Level 2 equations introduce a group-specific (or subject specific) intercept and slope. These two parameters come with

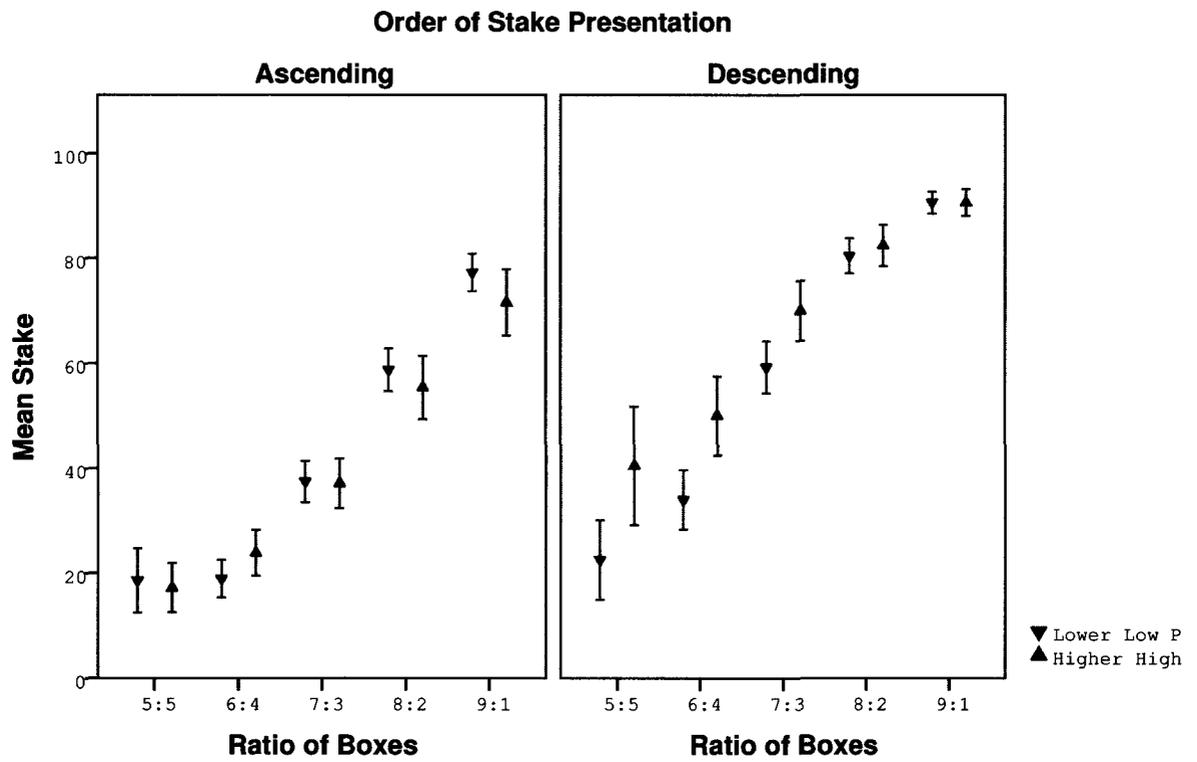


Figure 5.11: Money Staked by Box Ratio

their own residual terms σ_{int} and σ_{slp} , respectively. In effect, an LMM partitions the overall variance in a response variable into a deterministic part, which is predicted by a vector of X variables, and a stochastic part which consists of within-subjects error σ , and between groups variance in intercepts (σ_{int}) and slopes (σ_{slp}). The reason that LMMs are called “mixed” is that there is a mix of fixed effects given by the X’s and random effects coming from the group specific variance in the Z variables.

The model building approach involved adding candidate predictors for the dependent variable *Stake'*¹⁶ at each iteration, starting with a null model. The Akaike Information Criterion (AIC) was used to evaluate goodness of fit to ensure that the added predictors improved the model without sacrificing parsimony.¹⁷ As suggested by Figure 5.11, *Stake'* increases with *BoxRatio*, i.e. the greater the asymmetry, the higher the bet placed. The pattern of betting also appears to vary by *Order of Presentation* (Ascending vs. Descending), where HHP has a higher initial bet (intercept) in the Descending condition. *Box Ratio* is a fixed effect. A total of 5 models were tested using the R package `lme4` (Bates & Maechler, 2009; Ooms, 2009). These models are summarized in Table 5.11 and the code used to run the models is in Appendix K.

The final model in R code is:

$$Stake' = 1 + BoxRatio + Risktaking + Riskadj + (1 + BoxRatio|SubjType : Treat)$$

In words, this formula states that *Stake'* is a linear combination of the following

¹⁶The raw variable *Stake* was transformed using the Box-Cox method to make the distribution closer to multivariate normal, with

$$Stake' = Stake^{0.66}$$

¹⁷AIC behaves similarly to adjusted R squared in that predictors are added only if they improve the model’s predictive ability to a degree that is better than chance. AIC is in the form, “smaller is better”.

Table 5.11: Comparison of Linear Mixed Models for Amount Staked in the CGT

	Model 1	Model 2	Model 3	Model 4	Model 5			
					Estimate	SD	z value	p value
Fixed Effects								
Initial status (Intercept)	19.05	-15.57	-14.58	-30.38	-26.64	3.75	-7.09	< .001
BoxRatio	N/A	4.79	4.68	4.65	4.65	0.38	12.12	< .001
Risktaking	N/A	N/A	N/A	27.06	23.08	1.99	11.58	< .001
Risk adjustment	N/A	N/A	N/A	N/A	-0.67	0.23	-2.86	0.004
Variance Components								
Level 1: Within group (σ)	88.63	48.62	47.24	39.42	39.27	6.27		
Level 2: In initial status (σ_{int})	14.96	17.04	52.7	46.1	44.06	6.64		
In rate of change (σ_{stp})		0.00093	0.52	0.53	0.54	0.73		
Correlation of slope-intercept		-1	-0.89	-0.87	-0.86			
Akaike Information Criterion	13191	12119	12082	11757	11752			

Legend: Model 1: Unconditional Means Model, Model 2: Unconditional Growth Model;
 Model 3: Allow Random Slopes and Intercepts to vary per *SubjType* \times *Treat* subgroup;
 Model 4: Add *risktaking*;
 Model 5: Add *risk adjustment* (Final Model)

fixed effects: *BoxRatio*, *Risktaking*, and *Risk Adjustment*. The intercept (indicated by “1”) and the slope of *Stake* on *BoxRatio* are allowed to vary by combinations of *SubjectType* (HHP vs LLP) and *Treatment* (Ascending or Descending). Of the three fixed effects, two have a particularly strong positive association with amount staked: *BoxRatio* and *Risktaking*. As indicated in Table 5.11, *BoxRatio* and *Risktaking* are positively related to the amount staked. *Risk adjustment* is negatively related to amount staked as can be expected from the fact that LLPs scored higher compared to HHPs. The slope and intercept variance differs by *SubjectType* \times *Treatment* combination as can be visualized in Figure 5.12.

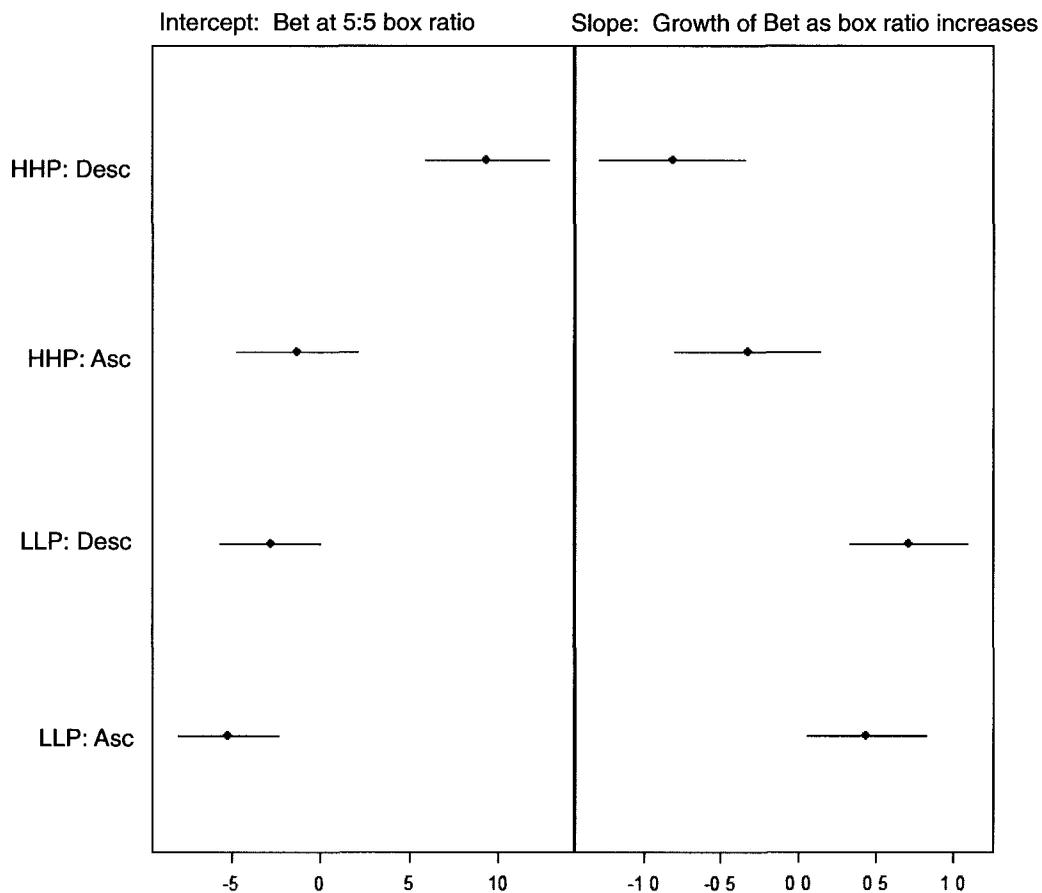


Figure 5.12: 95% Predicted Confidence Intervals of Intercept and Slope Variance

Figure 5.12 confirms the betting pattern in the overall sample analysis depicted in 5.11. What the restricted analysis adds is a model that fits the data. Also, the restricted analysis revealed that there is a difference not only in initial bets but also a correlation between the initial bet and rate of growth (slope) as *BoxRatio* increases. As Figure 5.12 shows, although HHPs start with a higher intercept in the *Descending* condition, their bets grow more slowly compared to LLPs. The LMM analysis also disentangled impulsivity from risk taking. In the model-building process, measures of impulsivity (*choice latency* and *response latency*) were entered into the model but did not improve goodness of fit, unlike *risktaking* and *risk adjustment*. Hence, HHPs placed higher initial bets as a result of higher *risktaking* and not impulsivity.

5.4.3 Section Summary and Comparison to Previous Research

The overall sample analysis and the restricted analysis both lead to the conclusion that more psychopathic subjects tend to take higher risks. This higher risk preference is independent from impulsivity or *delay aversion*. Two levels of analysis were used in the present investigation to amplify small effect sizes when considering the overall sample. This probably indicates that differences in risk preference have a high threshold of detection. Recall that the HHP subjects were in the 3rd percentile among 448 male students.

The result of the present study are consistent with a previous study (Deakin et al., 2004) in a normal population where a clear fractionation between *delay aversion* and *risktaking* was found. But contrary to that study, age and IQ¹⁸ were neither positively or negatively related to *risktaking*. The increased risktaking in the 5:5 box ratio by the more psychopathic subjects resembles the performance of inappropriate risk preference by alcoholics and problem gamblers (Lawrence et al., 2009) and in individ-

¹⁸In the present study, The Tower of Hanoi moves, time, and $\frac{\text{moves}}{\text{minute}}$ was a proxy for IQ

uals with prefrontal lesions (Clark et al., 2003). However, it is not clear that higher risk preference is a deficit in light of a recent study among entrepreneurs, who took greater risks compared to non-entrepreneurial managers. In that study, the authors hypothesized that this higher propensity for risk combined with cognitive flexibility underlie the innovation and creativity required to launch a new business (Lawrence et al., 2008). The strength of the present study is that points were converted to cash, and therefore risktaking was not a “theoretical” exercise. In both overall and restricted analysis, the more psychopathic subjects did not do worse than their less psychopathic counterparts.

Chapter 6

An Evolutionary Game Theoretic Model of the Prisoner Dilemma Results

Covenants struck without the sword are but words.

Thomas Hobbes, Leviathan

Until this chapter, the analysis of results has used standard statistical methods—e.g. regression, ANOVA, permutation tests and the like. Although valuable, they are quite limited in modeling evolutionary change because the findings are snapshots at a given time. Evolution, by definition, is a dynamic process in time, where mutations, no matter how slight, make a difference in the long run. Recall that in the Introduction the question was raised if there are functional reasons why psychopathy should exist. In Chapter 4, Hypothesis 4.3.1 stated: *High psychopathic participants paired against low will earn more than high psychopathic paired against high psychopathic individuals.* As stated in Chapter 5, one limitation of the human Prisoner Dilemma conducted is the very small sample of seven High P playing against seven Low P. In this chapter, the results of a virtual Prisoner Dilemma (i.e, a simulation using the actual participant moves) are reported.

In this chapter, two solution concepts from game theory are used: replicator

dynamics and evolutionarily stable strategies (ESS). The experimental results of the Prisoner Dilemma in the previous chapter are used in a virtual tournament and then findings are discussed in light of known game-theoretic results¹. Finally, I discuss how a High P strategy (characterized by high rate of defection) in the Prisoner Dilemma is a rational strategy using the Nash equilibrium concept.

A fundamental tenet in evolutionary game theoretic modeling is that entities (whether they are individuals, species, ideas, or groups) reproduce in proportion to the fitness points they get (Nowak, 2006; Maynard-Smith, 1982). The central actor is what is called a replicator, which is simply an agent that can make approximate copies of itself (Gintis, 2009). In the present study, monetary winnings are a proxy for fitness points. As mentioned in the introduction, game theory considers an action rational if it yields the most utility points. Evolution, which also applies to organisms that are beneath the threshold of rationality, uses the related concept of adaptiveness. In the Darwinian view, organisms that solve survival and reproductive problems produce more offspring (Tinbergen, 2005). True to its Darwinian roots, evolutionary psychology focuses on traits and behaviors (e.g. bonding, aggression, parenting) that are believed to increase reproductive success (Tooby & Cosmides, 2009; Daly & Wilson, 1988). Although it is not possible to conduct evolutionary experiments as such, there is value in modeling competition and the results thereof and then relating them to subject characteristics (High P/Low P).

We start by replicating Axelrod's computer tournament, but using our human subjects' strategies as the contestants.

¹The reason for focusing on the Prisoner Dilemma is that in game-theory, the IGT and CGT are games against Nature. Nobody wins against Nature.

6.1 A Virtual Prisoner Dilemma Tournament

In Axelrod's tournament, every strategy played every other strategy. Hence, a natural extension of the pairwise Prisoner Dilemma matches in Chapter 5 is to pair each strategy against every other strategy, including a clone of itself. For our 42 participants this means that there are 42×42 or 1,764 combinations of strategies. The strategies being considered here are the series of choices (of length 40) of **Cooperate** (C) or **Defect** (D) for each participant.

Three advantages can be gained from this simulation. First, since every strategy competes against the rest, a better picture can be gained of how good a particular strategy is—i.e. one strategy can be the best of 42 others, instead of being paired against a single opponent. Secondly, studying the relative ranking of the 42 strategies yields insight into whether High P strategies consistently outperform Low P ones.² Finally, one can determine whether in a mixed pair (High P vs Low P), there is indeed an advantage conferred to the High P strategy. Take note of an important limitation: strategies in this virtual tournament are “fixed” at the start and therefore do not adapt to the previous move of the other player. Hence, they are not very realistic. But since this limitation applies both to High P and Low P equally, then the virtual tournament is fair even if not entirely realistic. This is not a major hurdle since Axelrod's tournament players were programmed algorithms that were not meant to be prototypical human players. Hence, what is lost in flexibility is made up for by having strategies that were actually used by human players.

The procedure for this project's virtual tournament is as follows:

1. Match each strategy with every other strategy including its clone, yielding a 42×42 matrix.

²High P strategies are those used by High P participants. Low P strategies are those used by Low P participants. Each strategy was unique.

2. Parse each string of 40 choices letter by letter and do the same for its opponent. For each combination of choices (CC, CD, DC or DD), assign winnings according to the payoff matrix on Figure 4.1 on page 50.
3. Sum the winnings for each strategy.
4. Rank the strategies by points won, stratifying by *Subject Type* and test if one group performs better compared to the other.
5. Rank the strategies, this time stratifying by *Match Type* and conduct the statistical test.

The result of the virtual tournament is summarized in Table 6.1.

High P subjects won an average of 3,330.93, (SD = 637.38) while Low P subjects won an average of 3,128.50, (SD = 536.94). A non-significant t-test result of mean winnings showed that these did not vary by *Subject Type*, $t(40) = 1.12, p = .66$. Interestingly, as Table 6.1 shows, the top four strategies are taken by High P subjects, so a non-parametric test was conducted to test if one group ranked higher than the other. The Mann-Whitney test showed a non-significant result, $p = .10$, although as a group, Low P subjects had a higher mean rank of 19 as compared with 24 for High P subjects. Hence, the non-significant main effect of *Subject Type* is consistent with the Chapter 5 result.

In Chapter 5, it was reported that High P participants earned significantly more in the mixed condition as compared with playing against a High P opponent. Recall that in that analysis, the comparison consisted of seven mixed matches as against seven pure High P matches, making it a very small sample. In the virtual PD tournament, there was a chance to verify if the effect of *Match Type* remained when using 42×42 matches. Interestingly, the very same pattern held as summarized in Table 6.2. Table

Table 6.1: Virtual PD Tournament Winnings Ranking

Rank	Subject Type	Winnings	Rank	Subject Type	Winnings
1	High P	4,201.50	22	Low P	3,313.50
2	High P	4,197.00	23	High P	3,282.00
3	High P	4,078.50	24	High P	3,256.50
4	High P	3,918.00	25	High P	3,190.50
5	Low P	3,880.50	26	High P	3,175.50
6	High P	3,790.50	27	Low P	3,114.00
7	Low P	3,781.50	28	Low P	3,058.50
8	High P	3,709.50	29	Low P	2,989.50
9	High P	3,672.00	30	Low P	2,970.00
10	Low P	3,646.50	31	Low P	2,923.50
11	High P	3,567.00	32	High P	2,902.50
12	Low P	3,546.00	33	High P	2,860.50
13	High P	3,546.00	34	Low P	2,854.50
14	Low P	3,528.00	35	Low P	2,790.00
15	Low P	3,523.50	36	Low P	2,740.50
16	Low P	3,513.00	37	Low P	2,568.00
17	Low P	3,451.50	38	High P	2,566.50
18	Low P	3,448.50	39	Low P	2,382.00
19	High P	3,448.50	40	High P	2,256.00
20	High P	3,387.00	41	Low P	1,675.50
21	High P	3,324.00	42	High P	1,620.00

6.2 summarizes the mean amounts won by *Subject Type* and *Match Type*.

Table 6.2: Mean Winnings in the Virtual PD by Subject \times Match Type

Subject Type	MatchType	Mean Winnings	95% CI of the Mean
High	High-Low	84.55	81.49 to 87.61
Low	Low-Low	79.19	76.13 to 82.23
High	High-High	74.06	71.01 to 77.12
Low	Low-High	69.79	66.73 to 72.85

An ANOVA with winnings as the dependent variable and Bonferroni adjustment for multiple comparisons revealed the following results:

- Low P subjects paired with fellow Low P earned more than Low P's paired with a High P, $p < .001$, i.e. Line 2 vs. Line 4 of Table 6.2, (79.19 vs. 69.79).
- High P subjects paired with a Low P earned more than High P's paired with a fellow High P, $p < .001$, i.e. Line 1 vs. Line 3 of Table 6.2, (84.55 vs. 74.06).
- Among the mixed pairs, the Low P subjects earned less than their High P counterparts, $p < .001$, i.e. Line 4 vs. Line 1 of Table 6.2, (69.79 vs. 84.55).

Taken together, these results suggest that high psychopathy is not unconditionally advantageous, but dependent on the playing field and supports the argument for a frequency dependent nature of psychopathy as suggested by in Hypothesis 4.4.1 and by previous researchers (Mealey, 1995; Colman & Wilson, 1997).

6.2 Replicator Dynamics of the Virtual PD

In this section, a time element is introduced to the virtual PD model. Suppose the figures in Matrix 6.2 represent access to resources and mates. The replicator

dynamic equations essentially state, “to those who have more, more [offspring] will be given.” More precisely, replicator dynamics are described by the following system of differential equations (Nowak, 2006, p. 15).

$$\begin{aligned}\dot{x} &= x(a - \phi) \\ \dot{y} &= y(b - \phi)\end{aligned}\tag{6.1}$$

Equations 6.1 state that two competing species x and y (or types in our case) will grow or dwindle in the subsequent generation depending on how much they exceed or fall short of average fitness ϕ in the present. This is best understood together with the payoff matrix of the Prisoner Dilemma. With the 42×42 matched pairs in the virtual PD, we can summarize the average winnings when one *Subject Type* is matched against a similar or an opposite kind according to Matrix 6.2:

<i>Matrix 6.2</i>	Low P	High P
Low P	53.5	48
High P	68	47.5

Let x in Equations 6.1 stand for the frequency High P subjects, y the frequency of Low P subjects, and a and b for payoffs. \dot{x} and \dot{y} indicate the change in frequency of the types as a function of payoffs. Replicator dynamics analysis proceeds by identifying equilibria in order to predict long-run behavior. Equilibria are values of x and y such that when they are attained, the frequencies of x and y remain the same for subsequent time periods (i.e. both \dot{x} and \dot{y} remain 0). What are these values of x and y ? Matlab will be used to calculate the change in frequencies over time and find the equilibrium fraction of each type.

6.2.1 A Simulation of Replicator Dynamics in the virtual PD

One way of running the simulation is to take all 1,764 strategy pairs such as in Table 6.1. Two reasons why this is not very helpful are: (1) it is already known that the strategy with the highest earnings (i.e. Strategy # 1 of Table 6.1, which earned 4,201.50) will win against everyone else; (2) what is of interest are not the individual strategies themselves but the universe of High P and Low P strategies (i.e. in statistics they would be “random factors” and not “fixed factors”). How to select the strategy most typical of High P and of Low P subjects?

Hamming distance can be used as a measure to select the strategy that is “closest” to every other strategy in High P strategies, and the same of Low P strategies. Hamming distance is widely used in similar contexts, i.e., genetic relatedness (Pilcher, Wong, & Pillai, 2008) and character recognition (Bookstein, Klein, & Raita, 2001). It is computed by counting the the number of coordinates that need to be changed to make a string identical to another. For example, the Hamming distance between the strings 1011 and 1000 is two because it takes two mutations to change one to the other. In the present investigation, Hamming distance was computed using the R function `hamming.distance` (Dimitriadou et al., 2009) for each of the 21 High P and 21 Low P strategies. In Table 6.3 the strategies most similar to their own kind are listed. These are sequences of **Cooperate** (0) and **Defect** (1) moves over 40 rounds.

Table 6.3: Most ‘typical’ strategies

High P:	11101111011110111110110111111111111111
Low P:	11110111110110111011110111011101100111111101
<i>Legend: 0 = cooperate; 1 = defect</i>	

Note that both strategies have extremely high defection rates. The High P strategy

in particular is very close to “Always Defect” (all 1’s) except for five instances of cooperation. The Low P strategy likewise cooperated for only a total of nine instances.

These two strategies played against each other (and against a clone) and their frequency in the next generation was made proportional to the payoffs in Matrix 6.2. When paired against each other, the High P strategy earns 68 points and the Low P strategy earns 48 points. When paired against a clone of themselves, High P strategy earns 47.5 while the Low P strategy earns 53.5 points.³ The payoffs for encounters in one generation are summarized as follows.

$$\begin{array}{rcc}
 \textit{Matrix 6.2} & \text{Low P} & \text{High P} \\
 \text{Low P} & \left(\begin{array}{cc} 53.5 & 48 \end{array} \right) & \\
 \text{High P} & \left(\begin{array}{cc} 68 & 47.5 \end{array} \right) &
 \end{array}$$

How will a population of High P’s and Low P’s evolve over time? Assume that in the first generation only 1 percent of the general population are High P, in keeping with Neumann and Hare’s (2008) estimate of psychopathy in the general population. Figure 6.1 shows the the frequency of High P or Low P strategies over time. According to the model, by the 600th generation, the equilibrium frequencies of 3 percent for Low P and 97 percent for High P will have been achieved. In short, the replicator dynamic simulation predicts that a very small fraction (1 percent) of High P individuals will soon dominate the population.⁴

6.2.2 Possible Criticisms of the Replicator Dynamic Model

In this section, some possible criticisms of the model are anticipated and discussed.

³These points are the population average for the 42×42 individual round payoffs of (2.5, 2.5) if both cooperate, (1.0, 1.0) if both defect and so on.

⁴I gratefully acknowledge the Matlab code by Jack Hirschleifer and Juan Coll made publicly available at <http://www.socdynamics.org/id4.html>

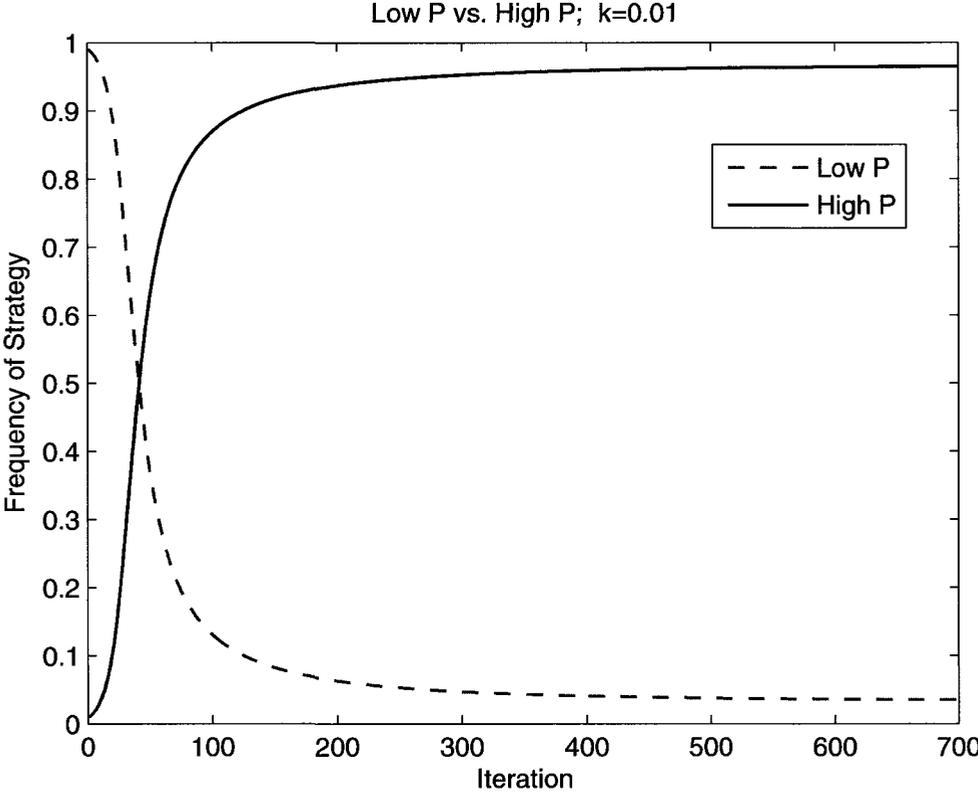


Figure 6.1: Evolutionary Course of High P, Low P Strategies

1. **Q:** In real life, psychopaths remain but a small fraction (1 percent by Hare and Neumann’s estimate); why is the replicator dynamic simulation “so far off the mark?”

A: The motivation behind the simulation is different from determining the actual incidence of psychopathy. The question being answered is the one posed at the **Introduction**: “are there functional reasons for why psychopathy should exist?” And the only outcomes being distinguished are whether: (a) one of the types goes extinct; or (b) both co-exist. For a more analytical treatment that enumerates the entire set of possible outcomes, see Appendix L. Note that the result of the simulation above corresponds to Scenario D.1 on page 189. In outcomes similar to Scenario D.1, High psychopathy is advantageous when paired against Low psychopathy but disadvantageous against its own kind. The relative proportions 97 percent and 3 percent are an artifact of the payoff values and may vary depending on what the true values are. Of note, it is a standard result for different types of individuals to co-exist in many simulations of the Prisoner Dilemma where defectors outnumber cooperators. See for example Christopher Hauert’s virtual prisoner dilemma lab at <http://www.univie.ac.at/virtuallabs/TwoByTwo/mixed.coexist.html>. The mix of psychopaths and non-psychopaths in the population can be thought of as a supply-and-demand problem: market price sets the equilibrium. In the replicator dynamic model, psychopathy’s payoff determines the optimal mix. Beyond a given frequency, psychopath numbers can only decrease because the net payoff to psychopathy decreases.⁵

2. **Q:** If the evolutionary dynamics are dependent on the payoff values, what if the

⁵I acknowledge Dr. Brook’s input to clarify this point.

payoffs are wrong?

A: Unlike in statistical models where parameter estimates require high precision, it is enough in replicator dynamics models for the payoffs to be ordinal in nature. This means that the payoffs should follow a Prisoner Dilemma structure, i.e.:

$$(HighP, LowP) > (LowP, LowP) > (HighP, HighP) > (LowP, HighP)$$

For as long as the order of these inequalities is followed, the evolutionary dynamics will be the same (Nowak, 2006, p. 75).⁶

3. **Q:** Many simulation models are sensitive to initial conditions (i.e. the initial frequencies of High P and Low P strategists.) How do we know that other initial values of the frequencies lead to the same result?

A: Provided only that each type is represented (i.e. not zero) in the population, different starting parameters will lead to the same equilibrium state. According to the equations 6.1, the proportion of types in the population changes according to how much each type exceeds or falls short of the overall average which is ϕ . We can set High P payoffs equal to Low P payoffs to find the steady state mixture. See Appendix H for a numerical calculation or Appendix I for a plot with different starting values of p that end in the same equilibrium frequencies of 3 percent and 97 percent.

6.3 Evolutionarily Stable Strategies

The real-world implication of the virtual PD simulation is that it pays to be

⁶The names of the choices here are arbitrary, they can just as well be (B,A), (A,A), (B,B), (A,B)

greedy, selfish, and opportunistic. The caveat is that if one happens to be paired with somebody of the same temperament, one ends up in a losing situation. One might speculate, quite pessimistically, that niceness evolved so as not to be stuck in mutual loss. Looking at the typical strings of High P and Low P, it is striking to note that they are not that different. Both types are more similar to “Always Defect” than to “Always Cooperate.” Does this mean that everyone is at heart opportunistic, and that it does not take a quantum leap from being a “normal person” to being a Levenson psychopath? ⁷ From the result of our experiments, it does appear that way from our sample of “normal” university students. To play “nice” by always cooperating invites exploitation. To play “nasty” by always defecting is likewise self-defeating because one is stuck in a bad outcome. This begs the question: how should any game ⁸ be played when it is not known that the other party is cooperative or greedy? One should try to play an “unbeatable strategy” known as an evolutionarily stable strategy (ESS) (?). First, a Matlab simulation will be done and then an analytical solution will be done.

It was assumed in the previous simulation that there are only two players (or strategies)—High P and Low P. In this next simulation, a third player, Tit-for-Tat (TFT) is added.⁹ Tit-for-Tat is chosen because it was the champion in Axelrod’s contests, and it is simple and easy to apply in real-life contests. The payoff matrix

⁷I am using this term to distinguish it with the Hare psychopath which involves an assessment of antisocial behavior from third parties. Levenson’s instrument on the other hand relies on first-person value judgments and expressions of beliefs regarding psychopathic traits.

⁸not necessarily the Prisoner Dilemma

⁹Strictly speaking, Tit-for-Tat is not a single strategy specified in advance (i.e. in terms of C’s and D’s) but one which adapts to its opponent’s last move. Hence, in this simulation there is a different Tit-for-Tat strategy for High P and Low P.

for these matches is as follows.¹⁰

<i>Matrix 6.3</i>	Low P	High P	TFT
Low P	53.5	48	69.5
High P	68	47.5	59
TFT	64.5	54	100

For the Matlab simulation, the initial player frequencies were set at 98 per cent for Low P and 1 per cent each for High P and Tit for Tat. The result of the simulation over 80 generations is depicted in Figure 6.2. Notice how different the outcome is of this match from the 2×2 case between High P and Low P. First, in this 3×3 case, both High P and Low P become extinct and TFT dominates the population. Secondly, notice how fast TFT gains ascendancy in this case: at generation 60, High P and Low P both go extinct. In the 2×2 simulation, neither strategy became extinct and it took about 600 generations for a stalemate to be achieved.

Why did the 3×3 simulation result in the extinction of TFT's opponents? TFT caused High P and Low P to go extinct because in the payoff Matrix 6.3, TFT is an "evolutionarily stable strategy" The precise formulation of an ESS is as follows:

$$E(A, A) > E(K, A)$$

$$\text{or if } E(A, A) = E(K, A)$$

$$E(A, K) > E(K, K) \text{ for all } K$$

(6.2)

¹⁰Note that the payoffs in between High P and Low P did not change from the 2×2 scenario. TFT payoffs with itself is 100 because TFT always cooperates in the first round and copies its opponent after that. But since it is playing itself then it gets 2.5 for each round for 40 rounds

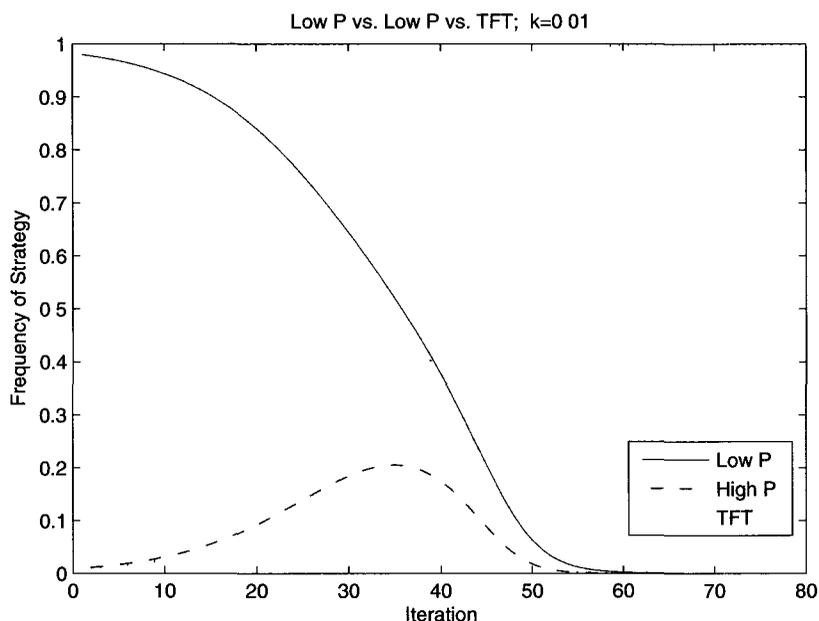


Figure 6.2: Evolutionary Course of High P, Low P Strategies, and TFT

In words: a strategy **A** is evolutionarily stable (i.e. its population share cannot be encroached upon by mutants) if either one of two conditions hold: (1) Any other strategy **K** pays less when matched against **A**. (2) If **A** and **K** pay equally well when playing against **A**, then **A** should pay more than **K** when paired against **K**. The phrase “for all **K**” is required for games with more than two strategies, to indicate that this condition should hold for every other strategy **K**.

To illustrate, Matrix 6.3 on page 106 will be used to identify the ESS. The procedure is very simple: look at the diagonal entries in the matrix. If a diagonal entry is the largest entry in its column, then the strategy to which it corresponds is an ESS.¹¹ If this strategy (call it **A**) is equal in value to another strategy, say it is **K**, then in column **K**, **A** should give a larger payoff than **K**. In Matrix 6.2 on page 101 neither High

¹¹For example, TFT is an ESS because it is the largest entry in its column (column 3). But neither High P nor Low P are ESS because their diagonal entries are not the largest in their respective columns.

P nor Low P are ESS's since the diagonal entries are not the largest values in their respective columns. As a result, the evolutionary dynamic is co-existence of the two strategies. ESS's can be generalized to not only Prisoner Dilemma type games but to all games. The so-called evolutionarily stable strategy would prevent alternative strategies (or mutants) from invading the incumbent population (?, ?, p. 15).

6.4 Psychopathy and Evolutionarily Stable Strategies

Now armed with the concept of an ESS, the question regarding psychopathy's existence becomes, "what are the conditions under which psychopathy should cease to exist?" The condition is easy to specify but hard to attain. Psychopathy as a life strategy will cease to exist only if some other pure strategy is an ESS.¹² It matters crucially whether the payoff structure is like the 2×2 matrix on page 101 or the 3×3 matrix on page 106. In the former, there is no pure strategy ESS whereas in the latter, TFT is a pure strategy ESS. Unfortunately, it has been shown that no pure strategy is evolutionarily stable (Boyd & Lorberbaum, 1987) and subsequently, that no strategy whatsoever whether pure or mixed (Lorberbaum, 1994) is evolutionarily stable in the repeated Prisoner Dilemma. The proofs given are mathematical and are beyond the scope of this project.

Dugatkin and Wilson (1991) describe how psychopathy would prosper in real-life Prisoner Dilemma situations. Whereas in laboratory matches, the pairing of players is maintained over a fixed number of rounds (i.e. 40 in our case), this restriction does not hold in real life. According to Dugatkin, in a population of individuals where

¹²Examples of pure strategies are "Always Defect" or "Always Cooperate". Mixed strategies on the other hand, involve choosing a move at random a *pre-determined* fraction. Tit-for-Tat is not a mixed strategy because its choice at a given move is fully determined by the opponent's previous move.

individuals are playing Prisoner Dilemma games in sub-groups, psychopaths would tend to rove from sub-group to sub-group in search of naïve cooperators. By playing “Defect” at every first turn and immediately finding a new partner, they accumulate winnings. Of necessity, the frequency of psychopathy would be limited by the number of cooperators and beyond a certain point, psychopathy (which we equate with pure defection) would be disadvantageous. This is where the importance of a psychopath’s ability to manipulate how he is perceived (i.e. being a con artist) becomes essential (Dugatkin, 1992).

In summary, assuming that the Prisoner Dilemma captured real-world encounters between psychopaths and non-psychopaths, the absence of an evolutionarily stable strategy rules out the extinction of psychopathy. The psychopath strategy would simply morph adaptively to take into account the rate of cooperation. Happily, the growth rate of psychopaths would be self-limiting. Unhappily for the rest of us, non-psychopaths inevitably would discover (whether by error or conscious intent), that defection pays against naïve cooperators. As such, the relative proportions of types would tend to oscillate back and forth until an equilibrium is reached.¹³

6.5 Psychopathy and Rationality

It was noted in the previous section that ESS’s exist infrequently, if at all, in the Prisoner Dilemma and perhaps so in other games. For example, in the well-known Rock-Paper-Scissors game, no pure strategy assures a win. How should one play games when no ESS’s are possible? Game theory has a slightly weaker prescription called Nash equilibrium and it is a criterion for rational play. In essence, a Nash equilibrium is a combination of moves such that each party is best responding to the

¹³Game theory goes so far as to prescribe that it would be rational for an individual to “mix” strategies, i.e. play defect sometimes and cooperate sometimes.

other party(-ies) moves. In this section, finding the Nash equilibrium in the Prisoner Dilemma will be described, and its relation to psychopathy will be discussed.

Consider once more the payoff structure of the PD as played by the subjects in the experiment. The italicized numbers (first in a pair) are row player payoffs and the other number is for the column player. How should a rational player play this game?

$$\begin{array}{cc} & \text{C} & \text{D} \\ \text{C} & (2.5, 2.5) & (0, 5) \\ \text{D} & (5, 0) & (\underline{1}, \underline{1}) \end{array}$$

Game theory prescribes that the row player reason as follows: if the column player plays C, then I get 2.5 if I play C and 5 if I play D. On the other hand, if the column player plays D, then I get 0 if I play C and 1 if I play D. Therefore, regardless of how the other party plays, I am better off playing D. The column player can use similar reasoning. If both acted upon their reasoning, they would end in cell (D,D) with the payoffs underlined in the payoff matrix. This outcome, which is reached when each agent considers only the best for oneself¹⁴, is known as a Nash equilibrium. If the game is not repeated, i.e. it is a single round, then Defection is the *only* Nash equilibrium, which means it is the only rational course of action. To do otherwise would be irrational by the criterion of acting on behalf of one's self-interest.

In the repeated game however, such as the one played by our human subjects, there is mutual benefit to be attained in both players cooperating. This does not mean that playing "Always Defect" is not a Nash equilibrium—for it is. It is just that in the repeated game, there are *multiple* Nash equilibria, of which Tit-for-Tat

¹⁴This is an important point, since if both agents thought of their common interest, they would choose C,C.

is but one (Heap & Varoufakis, 2004). But the crucial point is that one cannot rule out “Always Defect” (i.e. the closest in Hamming Distance to the High P strategy) as being irrational. Why so? It is because a choice C has a chance of being exploited by a move of D by the other player resulting in the lowest payoff of 0. With every choice D on the other hand, one is “immunized” from exploitation and hence, one is assured of never getting the lowest payoff. In this sense, “Always Defect” is a rational strategy because the worst payoff one can get is 1 in every round. In game-theoretic language, “Always Defect” maximizes the minimum amount that one can receive.

6.6 Chapter Summary

In this chapter, the raw choices of 42 human subjects in the PD were entered into a virtual competition. The outcomes of each 40-round encounter with every other strategy and a clone revealed a non-significant *Subject Type* effect on *Winnings* and a significant *Subject Type* \times *Match Type* effect. These are consistent with Chapter 5 results. Subsequently, a “most typical” strategy for each *Subject Type* was chosen using Hamming distance as a criterion and these two representatives were made to compete. Population share in subsequent generations was modeled using the replicator dynamics equation. The dynamics showed that High P (which defects more) beats Low P (which defects less). This outcome does not reflect what happens in real life because: (1) players are free to choose other partners; (2) not all real-life encounters are Prisoner-Dilemma like in structure (e.g. Rock-Paper-Scissors); and (3) mixing of strategies happens. Nevertheless, if the Prisoner Dilemma modeled at least *some* human interactions, the High P strategy of defecting more often (or all the time) is a rational strategy because it is a Nash equilibrium.

Chapter 7

Discussion

God knows Judas not pay.

Filipino sticker in public transport.

Robinson Crusoe spent 28 years in a remote island before returning to civilization. Similarly, the dominant metaphor for the mind in the last few decades has been the computer. For many research programs in psychology, the social milieu of the mind can be safely bracketed aside. However, for a phenomenon as complex as psychopathy, the computer metaphor probably needs to be stretched in light of the fact that nowadays, no computer is an island, but a node in a rich network of connections. For sure, the deficits observed in psychopathy are no less true today than in the time of Cleckley. Psychopaths have deficient fear conditioning, shallow affect, impaired somatic markers, high risk preference, and the like. But at least some of these very same characteristics, when embedded in social interactions, turn into advantages. In fact, some researchers have argued that to conceive of psychopathic traits as deficits is misguided:

It is difficult for us to conceive of a neurocognitive defect that could enhance such qualities as lying, conning, manipulation, glibness, and charm. It seems a logical contradiction to suppose that a disorder could improve cognitive abilities. (Harris, Skilling, & Rice, 2001, p. 230)

Although the “deficits paradigm” is still the dominant school of thought, evidence

is accumulating that some of these deficits are part of a life strategy that is adaptive. Briefly, psychopathy is adaptive because psychopaths are shrewd sexual strategists, can spot the most vulnerable victims, and perform competently in laboratory experiments. Additionally, healthy individuals periodically engage in violence, lying, and deception, and this behavior is accepted in varying degrees by society's norms, if not by its laws. It is time to revisit the research questions in the Introduction and integrate with them the results of the present investigation.

7.1 Summary of Experimental Results

1. **On selfish strategies in the Prisoner Dilemma.** Mixed matches vis-à-vis homogeneous matches are profitable for high psychopathic individuals. This finding was in line with Hypothesis 4.3.1 which stated: *High psychopathic participants paired against low will earn more than high psychopathic paired against high psychopathic ones.* This result is consistent with three previous studies (Mokros et al., 2008; Rilling et al., 2007; Montañes et al., 2003) but inconsistent with those of Widom (1976). Whereas high psychopathic individuals thrive in mixed matches, they also earn significantly less than low psychopathic individuals in homogeneous matches. This divergent outcome depending on match type is a result of greater defection rates on the part of high psychopathic individuals. These results held true both in the human matches and in the simulated Prisoner Dilemma where each strategy was made to play against every other strategy. A counter-intuitive finding that was obtained in the human matches is that rate of cooperation did not vary significantly between subject types as measured by *Self-advantageous Defection Index* or the ratio of {D,C} to {C,D} moves. This result suggests that the difference in propensity to defect is a mat-

ter of degree and that it is possible for one type to behave as if they were the other. However, this small degree of difference in rate of cooperation is enough to make a huge difference in mixed matches in favor of high psychopathic individuals. Therefore, high psychopathic individuals are more selfish than low psychopathic individuals *only* in mixed matches, but in homogeneous matches, subject types are not significantly different.

2. **On learning from punishment and response reversal.** The Iowa Gambling Task revealed no significant difference in number of picks from good decks, NET scores, and good deck streakiness. Hence, no evidence was found to support Hypothesis 4.3.2: *High psychopathic participants will have a lower NET score in the IGT compared to low psychopathic ones.* Both subject types learned to favor the good decks by the 20th move. Importantly, the performance measures of subjects in this experiment are favorable to or comparable with population norms. The null finding in the present study is consistent with the results of Schmitt et al. (1999) but inconsistent with three other studies (Honk et al., 2002; Mitchell et al., 2002; Blair et al., 2001). The nonsignificant difference in good deck streakiness needs to be tempered by the result that high psychopathic subjects were less streaky for bad decks ostensibly as a result of more bad deck losses sustained by high psychopathic subjects. Hence, contrary to mainstream literature in the Iowa Gambling Task, not all psychopathic subjects are afflicted with response reversal and passive avoidance deficits—deficits that are found in substance dependent and ventromedial prefrontal cortex patients.
3. **On higher preference for risk and impulsivity.** The Cambridge Gamble Task result disentangled propensity for risk from impulsivity (delay aversion). In the overall sample analysis, while high psychopathic subjects had longer la-

tency times for choice of color and amount bet than low psychopathic subjects, the former tended to place higher bets when the chance for a win was *merely* 50 percent. This result in the overall sample was confirmed in the subsample consisting only of the most and least psychopathic extremes. The strongest predictors for (a) differences in initial bets (i.e. at the 5:5 box ratio) and (b) the growth of bets over different box ratios were risk adjustment and risk preference. Specifically, low psychopathic subjects tended to align their bets with chances of winning and high psychopathic subjects tended to place higher bets when the odds of winning were roughly even. Importantly, however, the overall monetary winnings of both subject types were the same, except in the Descending condition, where high psychopathic subjects had marginally significant ($p = .08$) higher wins. Thus we did not find evidence to support Hypothesis 4.3.3: *High psychopathic participants will earn more money in the CGT compared to low psychopathic ones.* The high psychopathic individuals in our experiment seemed to resemble the Silicon Fen entrepreneurs (Lawrence et al., 2008) in their appetite for risk although the former did not outperform their low psychopathic peers.

Overall, it was only in the Prisoner Dilemma task that high psychopathic subjects had an advantage, and only when paired against a low psychopathic subject. The overall monetary winnings for both groups were about even. This null result in overall monetary winnings suggests that the theorized deficits in risktaking and passive avoidance are not enough to undermine the advantage that high psychopathic subjects have in social interactions as modeled by the Prisoner Dilemma. There are two possible interpretations to the profile of results obtained.

First, it could be that the subject types in our sample, despite belonging to ex-

treme quartiles, were not substantially different with respect to psychopathic traits. The main argument for this view is that our “psychopaths” were obtained from a university population and most probably did not have the antisocial tendencies of psychopaths diagnosed through the PCL–R. Factors such as having a good education, equivalent cognitive abilities as measured in the Tower of Hanoi favor the view that our high psychopathic subjects resemble the Silicon Fen entrepreneurs whose higher appetite for risk also made them more innovative businessmen. The main weakness of this perspective is that even if our “psychopaths” were not “psychopathic enough”, there was still a major advantage in favor of high psychopathic ones in mixed matches in the Prisoner Dilemma. If Hare’s estimate is true that psychopaths constitute only one percent of the general population, then most matches that psychopaths will find themselves in would be of the mixed type. Hence an alternative interpretation of the results is worth considering.

Second, it could be that among successful (i.e. well-adjusted) psychopaths, the threshold for self-destructive behavior is higher than the threshold for exploitative ability in social dilemmas. It might be that successful psychopaths share some of the risktaking traits of PCL-R diagnosed psychopaths, but whatever deficits in this area there are, these are more than offset by the rewards brought about by a higher propensity for defection. According to Kiyonari et al. (2000), individuals have a social heuristic to cooperate in the one-round Prisoner Dilemma, even if the rational choice is defection. The result of the present investigation suggests that this social heuristic, if present, is not present to the same degree in everyone. High psychopathic subjects find the monetary reward more compelling than the positive emotions evoked by mutual cooperation. Psychological satisfaction is good, but money is even better. In light of the results obtained in this study, the following operational definition of a successful psychopath is offered:

A “successful” psychopath is one:

1. whose propensity for acting on behalf of one’s self-interest exceeds the propensity to act for the common good;
2. whose social savvy brings about rewards that outweigh deficits in the sub-social (i.e. personal) realm.

Stated differently, a psychopath is “successful” if it is easier for him to win on account of social savvy than to lose on account of personal impairments.

7.2 Psychopaths and Non-psychopaths: Cut from the Same Cloth

For me, the most perplexing result in the Prisoner Dilemma is not that psychopaths outdid non-psychopaths in the mixed pair matches. It was this: why was *Self-advantageous Defection Index* not different between subject types? How to reconcile this outcome with Kiyonari et al.’s claim that humans have a social heuristic to cooperate? Had low psychopathic subjects cooperated in homogeneous matches, each party in a pair would have maximized his winnings. My hypothesis is that there is a selfish drive in each individual regardless of psychopathy status. In situations where there is a choice between fairness and rationality (i.e. selfishness), individuals might opt for one or the other. It is known in Prisoner Dilemma games that there is no clear “cooperator” and “defector” personality types (Frank, 1996). This fact reveals more about human nature in general than about psychopaths in particular. The rationale for the similarity of Prisoner Dilemma choices (i.e. the strings of length 40) in Chapter 6 can be explained by revisiting the concept of mixed strategy.

Colman and Wilson (1997) suggested an evolutionary game-theoretic model that

supports why antisocial personality persists at a small frequency.¹ Other researchers make the same claim of psychopathy (Barr & Quinsey, 2004; Glenn & Raine, 2009). In brief, their argument is that psychopathy (or APD) persists at an equilibrium frequency k where the weighted payoff of a selfish action is equal to the weighted payoff of a cooperative action. To clarify the analysis, the concept of mixed strategy is now discussed. As mentioned in Chapter 6, a mixed strategy involves randomizing between two or more choices (or actions) for a predetermined fraction of the whole. There is a need to distinguish two senses of *fraction* here:

1. *Fraction of the population.* In this view, a certain percentage of the population plays a pure strategy. For example, in the game Rock, Paper, Scissors this interpretation means that there are three subtypes in the population, with each one playing a single strategy all the time.
2. *Fraction of the time.* In this view, *everyone* in the population plays each strategy a certain fraction of the time, i.e., play each strategy (R, P, S) one-third of the time. This is the typical way that Rock, Paper, Scissors is played.

Colman and Wilson (1997) favor the first sense of mixed strategy [*emphasis mine*]:

There are two possible interpretations of the mixed-strategy equilibrium in an evolutionary game of this type. Either the population evolves to a dimorphic mixture of cooperative and antisocial phenotypes, in the proportions specified by the parameters of the model, in this case 98 percent and 2 percent respectively, or it evolves to a form in which every individual shows a propensity to randomize between behaving cooperatively and antisocially in the required proportions. *The first interpretation is clearly more likely to be correct in this case . . .* (p. 30)

However, there are important reasons why the second sense of mixed strategy (i.e. fraction of the time) is more realistic. First, by alternating between strategies, one gets

¹Strictly speaking, their research has nothing to do with psychopaths. I am citing their study here because the present study is an empirical test of their model, had they modeled psychopath behavior in the first place.

unpredictability, which in turn foils and discourages opportunistic play. For instance in penalty kicks at soccer games, players are known to employ a mixed strategy (dominant foot vs. non-dominant foot) to increase chances of scoring (Chiappori, Levitt, & Groseclose, 2002). Secondly, the alternation of strategies allows for reciprocity to arise. For example, in politics, there is the dynamic of “wheeling-and-dealing” where the content of a particular legislation is not the sole criterion, but also future support when the time comes for a legislator to lobby for his own law. Clearly, the Prisoner Dilemma results in this investigation—where rate of cooperation did not vary significantly between subject types—favor the second sense of mixed strategy. Still, it is useful to think of cooperators and defectors although there is a range of intermediates between one and the other. The issue of detecting types (or overall tendencies) when faced with scenarios that allow for cooperation to emerge becomes important.

7.3 Honesty/Dishonesty Arms Race

Suppose that in Prisoner Dilemma-like situations, there is a mechanism for cooperators to find fellow cooperators. Richard Dawkins (1976) called such a mechanism “green beard”. Theoretically, carriers of such an allele would pass it on to progeny so that kindred types would be able to recognize each other. At the same time, its absence in another individual would warn a green beard individual not act cooperatively. In game theory, signalling games refer to those in which an individual receives “messages” but is unsure of the “type” of the individual sending the message.

A Prisoner Dilemma experiment in which subjects were allowed to communicate with their partner prior to a one-shot game revealed that individuals were able to predict “fellow cooperators” at a rate better than chance (Frank, 1996). An important feature of that experiment was that partners could see each other face-to-face, thus

allowing signals such as tone of voice, eye contact, speech and the like to serve as indicators of honesty. Imagine for example that in a Prisoner Dilemma situation, you could be paired with a 100 percent honest individual who said, “I will play cooperate.” In that case, I would always play “Cooperate” and both of us would make the maximum winnable amount. In another Prisoner Dilemma experiment, Kiyonari et al. (2000) had the participants make their choice *sequentially* instead of *simultaneously*. First, one partner in a pair made a move then revealed the choice he/she had made. The second partner was then given the chance to make a move. Surprisingly, a high number of second players played “Cooperate” even when the rational move was to defect, not only because it paid more, but also because no retaliation in the future was possible.

If one had to distill the the essence of the Prisoner Dilemma in two words, I would choose, “coordination” and “commitment”. The problem of coordination is ensuring that one lands in the {Cooperate. Cooperate} outcome.² The problem of “commitment” arises because one is vulnerable to exploitation by a selfish partner. In real-life situations, one has little to go by except for a handshake, a verbal agreement, a certain look of the eye, all of which are fragile guarantees. Enforcing commitment between individuals is a subject of philosophical treatises. Thomas Hobbes (1651) [1651], a philosopher who had a pessimistic view of human nature, conceived of the State as the arbiter and guarantor of contracts between individuals, who could not be left to their own devices. By contrast, some theorists (Hirshleifer, 2001; Frank, 1996; Nesse, 1990) argued that emotions serve the function of regulating cooperation and reciprocity. For Frank (1996), sympathy is a commitment device that allows cooperation to emerge:

²At least for non-psychopaths this would be the case. I allow for the possibility that psychopaths might have the goal of winning at the expense of the other, but then they would still need to coordinate their choices with the cooperative choice of a sucker.

The emergence of sympathetic bonds among people is a very complex, physical, cognitive, and emotional dance. People feel one another out, respond to one another, choose to develop closer bonds with some, and abandon further contact with others. (Frank, 1996, p. 18)

Selfish individuals on the other hand, are aware of the value of emotions and can react in two ways: (1) mimicry, and (2) spotting the most vulnerable victims. The latter was dealt with in Chapter 3 regarding the mysterious powers of psychopaths who can spot the easiest victims by their gait alone. The former has been known for quite sometime and Machiavelli (1532) [1532] is perhaps its most famous proponent:

Therefore it is unnecessary for a prince to have all the good qualities I have enumerated, but it is very necessary to appear to have them. And I shall dare say this also, that to have them and always to observe them is injurious, and to appear to have them is useful; to appear merciful, faithful, humane, religious, upright, and to be so, but with a mind so framed that should you require not to be so, you may be able and know how to change to the opposite. (Machiavelli, 1532, p. 95)

In summary, the need for coordination and commitments in Prisoner Dilemma-like situations probably gave rise to emotions as important signals to distinguish friend from foe. To some extent, cooperators can find each other, but selfish individuals are determined to simulate the appearance of genuine emotion. The conflicting aims of the two types becomes an evolutionary arms race. Up to now, I have discussed the external battle between psychopaths and non-psychopaths. In the next section, I will discuss the war within the mind.

7.4 The Psychopath's Cognitive Stance

Since the present investigation is a project in cognitive science, it would be incomplete without discussing (quite speculatively) what the mind of a psychopath might be like. In a word, I would guess it would be this: fragmented. The mental stance

of a psychopath would be: to act in pursuit of selfish motives and yet, being dependent on the goodwill of others, to appear truly altruistic. Since I have argued that healthy individuals also engage in psychopathic tactics, then I take the psychopathic stance to apply to everyone in varying degrees. Cleckley was prescient in illustrating the cognitive dissonance as a “madness” afflicting the characters he referred to as incomplete manifestations. I have entitled this project *Sons of Barabbas or Sons of Descartes* not so much to argue that the psychopath is one or the other but to suggest that there are different types—not only the Hannibal Lecter types. Every age has its signature psychopathic crime. If in 19th century it was murder (e.g., Jack the Ripper), in the 20th century, genocide (e.g., Hitler), then I suggest that these days, it must be greed (e.g., Wall Street). What cognitive mechanisms might exist to ease the discomfort of cognitive dissonance—at least for those of us who have a conscience? How does one uphold the unity of the Self in the face of a glaring disparity between words and actions, or between conflicting desires?

Trivers (2002) suggested self-deception as a mechanism. The best liar, according to Trivers, is somebody who believes the lie himself. In this way, the mind does not need to keep the story straight. Lenin is supposed to have said, “A lie repeated often enough becomes a truth.” Self-deception in itself is not malevolent. Each person is “self-deceived” in the sense that one does not perceive the blindspot (scotoma) that is caused by the passage of the optic nerve in front of the retina. The brain simply “fills in the gap”. In Chapter 3, two other cognitive mechanisms were suggested: i.e. euphemistic language (Smith, 2004) and ethical fading (Tenbrunsel & Messick, 2004) in business dealings. In working with inmates in a psychiatric hospital, I learned a very commonly used technique: minimization of blame. This tactic involves playing down one's role, or better yet shifting the blame to the victim. Another convenient kind of self-deception is when the victim is faceless, if there is a victim at all. Guilty

of this kind of deception is Marc Hauser, and in his case, it might be called *hubris*. In classical literature, this sin involves accomplishing something extraordinary and as a result thereof, losing touch with reality. Fabrication of data (Wade, 2010) might appear to be a contemporary phenomenon, but the basic plot of Hauser's story is like the scholar Faust's deal with the devil by Göethe (1853). To do an evil act, one does not need to be extraordinarily bad, it is enough to be ordinary. In the words of the philosopher Paul Ricoeur:

To say that man is fallible is to say that the limitation peculiar to a being who does not coincide with himself is the primordial weakness from which evil arises. (Ricoeur, 1986, p. 146)

I take “not coinciding with himself” as the predicament of wanting to do the good but lacking the will to do so. Consider a professional executioner who has the terrible task of ending some criminal's life. The justification could be something like, “I have to have this job to feed my family” or “If I did not do it, somebody else would.” In summary, sophisticated mechanisms exist in the brain that ease cognitive dissonance that would otherwise cause great anxiety. Psychopaths, if they have no conscience, would need less help from these mechanisms. But ordinary mortals—the man on the street, the cashier at the grocery—possess these very same mechanisms. So, each individual is capable of acts normally attributed to psychopaths.

7.5 Limitations and Future Research

The foremost limitation in the present investigation lies in its interdisciplinary nature. For a dissertation in a specialized field to be right, it is enough that that the science in that field works. An interdisciplinary dissertation such as the present work depends on the correctness of the science of each respective area. I have tried to bridge findings in game theory, evolution, and psychology in an attempt to explain

psychopathy. Psychology is a more or less established field, but evolution and game theory, and particularly, the mathematical approaches involved in them are still in a state of flux.

In terms of experimental design, the comparison of psychopathic and non-psychopathic individuals would have been greatly improved with the availability of physiological measures of fear and risktaking. For example, were the skin conductance responses different between subject types different when picking from bad decks in the Iowa Gambling Task? Regarding bad deck streakiness in the IGT, it was mentioned that high psychopathic subjects had a higher rate of losses from bad decks. This is a purely artifactual outcome coming as a result of the stochastic nature of losses in the task. It would have been interesting to find out whether a greater rate of losing in the high psychopathic group made a difference in terms of skin conductance measures when compared to the low psychopathic group. A similar physiological measure in the Cambridge Gamble Task would have been very useful. In particular, since one of the CGT's objectives is to study "hot decision-making", it would have been extremely helpful to have the physiological measures of high psychopathic individuals when placing a high bet. Was such betting "cold-blooded" or accompanied with some trepidation?

In the Prisoner Dilemma, the finding that high psychopathic individuals earn more than low psychopathic individuals in mixed matches used a liberal p value of .055 as a basis of significance. Thus, a bigger sample size might have resulted in more statistical power. More importantly, self-reports of (i) satisfaction over the earnings, and (ii) satisfaction with the cooperativeness of the partner might have given a window into the emotions accompanying outcomes. Nesse (1990) for example, reported about feelings involved in the Prisoner Dilemma. Also a self-report measure would have given insight into whether the social heuristic of cooperation exists to the

same degree between psychopathic and non-psychopathic subjects.

On the positive side, the greatest strength of the experimental set-up was the double-blinding in subjects and experiment administrator. On the one hand, the participants had no idea that the specific traits being studied were psychopathic traits. On the other hand, the experiment administrator³ did not know the Levenson profiles of the participants. Also, the use of money as an incentive probably guarded against “fake responses.” In hindsight, a logistical improvement might have been made if money was given as a reward after each trial, as opposed to a lump sum at the end of all three games. It is not known if larger amounts of money involved would heighten the difference between psychopathic and non-psychopathic individuals, although that seems a reasonable assumption to make—except for whoever ends up paying for the bill.

The choice of a modeling tool (i.e. game-theory) was motivated by self-preference and not by necessity. As a prescriptive discipline that proposes what an ideal rational actor should do, game theory does not always describe human and animal behavior accurately. For example, motorists leave tips for waitresses in roadside cafes even if the same motorists know that they are unlikely to receive something in return for their generosity (Frank, 1996). This behavior is inconsistent with the self-interest maxim that is the cornerstone of game theory. Monkeys who were hitherto happy to exchange a token with a cucumber became furious and refused to exchange once they observed that a conspecific received grapes (a more valued reward) for the same task (Brosnan & De Waal, 2003). If monkeys acted on the basis of self-interest alone, it would be more “rational” to accept a slice of cucumber instead of receiving nothing. In short, game theory falls short as a theory of human or animal behavior. Behavior can appear to be motivated by concerns over equity and cooperation (Mailath, 1998).

³i.e. Sean Leach

Even worse, game theory does not take into account the cognitive limitations on the part of human game players in prescribing equilibrium solutions. Two such limitations are : (i) most humans lack the ability to calculate optimal probability weights for each action in a repeated game; and (ii) humans are averse to or incapable of randomizing choices in accordance with such probabilities (West, Lebiere, & Bothell, 2006). Erev and Roth (1998) who compared three kinds of models to account for human game playing reported that random guessing outperforms the game theoretic prescription in five of twelve games. Because repeated games can have enormous strategy sets which humans are incapable of remembering, Erev and Roth proposed an adaptive learning approach that is more human cognition-friendly and is based on a few rules. Some of these rules are: (i) select those choices that did well in the past, (ii) update choice probabilities by large increments at the start, but only in little increments as the game progresses; (iii) recent choices should be given more weight than older choices. In their meta-analysis of game playing, Erev and Roth's adaptive learning approach came out as the best model, beating random guessing and the game-theoretic equilibrium prescription.

If game theory yields limited insight into how humans play games, it nevertheless serves as an excellent rationale for why psychopathy exists. Psychopathy exists because it is compatible with the same self-interest maximization principle of game theory. More to the point: psychopaths exist because as the result of the PD task shows, their strategies better conform to game-theoretic prescriptions compared with non-psychopaths. In itself, game theory is agnostic about the cognitive architecture of players—the canonical solutions it offers for the Prisoner Dilemma are the same whether the player is human or machine. The divide between human game playing and game theoretic prescription is partly due to the false choice between selecting an abstract rational actor (whose abilities might exceed human players, but always

find the equilibrium solution) on the one hand, and selecting a learning model (which might be human-like but be slow to converge upon the equilibrium solution) on the other. But these choices are not mutually exclusive, allowing for collaboration between the two camps as exemplified in the work of West and Lebiere (West, Stewart, Lebiere, & Chandrasekharan, n.d.; Lebiere & West, 1999).

Research in the interface of games theory with other disciplines is quite young and there are emerging areas of inquiry variously known as experimental economics, behavioral economics, social neuroscience, and neuroeconomics. Instead of suggesting newfangled areas of research collaboration, which is of course fruitful, I would like to recommend, quite ironically, that professional philosophers should sail into these new frontiers. To cite a concrete example, Glenn and Raine (2009) discuss how the use of fMRI studies in conjunction with Prisoner Dilemma games sheds light on *where* morality is in the brain. Quite unfortunately, the very same authors refer to psychopaths as the “impersonation of evil”. My issue is that there seems to be a mistaken mapping between cooperation as “the good choice” and defection as “the bad choice” especially in light of the fact that “good feelings” result from cooperation and “feelings of betrayal” come from defection. In other papers (Boyd, 1988; Ellison, 1994), the defectors are labeled “cheaters”. From a philosophical standpoint, two serious objections can be made for the sake of conceptual clarification.

First, the mapping of good with cooperation and bad with defection is problematic. To see this, consider again the issue of Hauser’s graduate students speaking up about suspicions of data fabrication. The game-theoretic representation of this scenario would be {Cooperate, Cooperate} meaning “keep quiet” on the part of both Hauser and the students. That outcome would have been advantageous to both the careers of Hauser and the students (but bad for science). This shows that a game-theoretic analysis, and even one supplemented with high technology does not define

what morality is. Morality is the philosopher's cup of tea, and their contribution to modern-day questions of ethics is most valuable. A paradigmatic example of neuroethical research is Greene et al.'s (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001) trolley and drawbridge problems. In the trolley problem, a runaway trolley is heading towards five people, but an observer can flip a switch to redirect the trolley to a separate track where only one person will be killed. Should the observer flip the switch? Most people answer yes. In a related problem, the observer is on a drawbridge above the track together with a burly man. Should the observer push the burly man to stop the trolley? Most people say no. According to Greene, the only difference between these two scenarios is that the latter is more emotionally salient compared to the first. So the philosophical question is: are emotions relevant to moral judgments? Kant would probably not be too happy.

Secondly, on the issue of defection construed as "cheating", a problematic assumption is made. This assumption is: *Defection is against the rules*. But I am not aware of a Prisoner Dilemma experiment that explicitly forbids defection. Rather, what seems to be happening is that a social norm for cooperation influences the how the dilemma is expected to be resolved. Kiyonari et al. (2000) suggest that most participants in the Prisoner Dilemma (especially those who expect cooperation from their partners) undergo a subjective transformation of the Prisoner Dilemma game into an Assurance game. See pages 413-414 of Kiyonari et al. (2000) for details. On the other hand, the game-theoretic prescription (at least for a single stage Prisoner Dilemma) is to defect because it strictly dominates cooperate (i.e., defection's payoffs are superior regardless of what the opponent does). In effect, there is a clash of terminologies reflecting a clash of values—should one, in playing "Defect", be rational at the cost of being a "cheater", or should one, in playing "Cooperate", be "nice" and commit a logical mistake? This last point reveals why the Prisoner Dilemma has

176,000 results in Google Scholar and why it might offer a clue as to why psychopathy exists in the population.

Bibliography

- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, 4(3), 165–178.
- Albert, J. (2008). Streaky hitting in baseball. *Journal of Quantitative Analysis in Sports*, 4(1), 3.
- American Psychiatric Association. (n.d.). *DSM-5 Development*. Retrieved August 27, 2010, from DSM-5: <http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=16#>.
- American Psychiatric Association. (1968). *Statistical Manual Mental Disorders*. Washington, D.C.: American Psychiatric Association.
- American Psychiatric Association. (1980). *DSM-III*. Washington, D.C: American Psychiatric Association.
- Armstrong, D. (2008). *Doctor didn't disclose glaxo payments, senator says*. <http://online.wsj.com/article/SB122304669813202429.html>.
- Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science*, 211(4489), 1390.
- Baron-Cohen, S., Leslie, A., Frith, U., et al. (1985). Does the autistic child have a “theory of mind”. *Cognition*, 21(1), 37–46.
- Barr, K., & Quinsey, V. (2004). Is psychopathy a pathology or a life strategy? Implications for social policy. *Evolutionary psychology, public policy, and personal*

- decisions*, 293–317.
- Bates, D., & Maechler, M. (2009). *lme4: Linear mixed-effects models using s4 classes* [Computer software manual]. Available from <http://CRAN.R-project.org/package=lme4> (R package version 0.999375-31)
- Bazerman, M., Loewenstein, G., & Moore, D. (2002). Why good accountants do bad audits. *Harvard Business Review*, *80*(11), 96–103.
- Beauregard, M., Chertkow, H., Bub, D., Murtha, S., Dixon, R., & Evans, A. (1997). The neural substrate for concrete, abstract, and emotional word lexica: A positron emission tomography study. *Journal of Cognitive Neuroscience*, *9*(4), 441–461.
- Bechara, A. (2007). Iowa Gambling Task professional manual. *Boca Raton, FL: Psychological Assessment Resources*.
- Bechara, A., Damasio, H., Tranel, D., & Anderson, S. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience*, *18*(1), 428.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293.
- Bechara, A., Dolan, S., & Hindes, A. (2002). Decision-making and addiction (part II): Myopia for the future or hypersensitivity to reward? *Neuropsychologia*, *40*(10), 1690–1705.
- Benning, S., Patrick, C., Blonigen, D., Hicks, B., & Iacono, W. (2005). Estimating facets of psychopathy from normal personality traits: A step toward community epidemiological investigations. *Assessment*, *12*(1), 3.
- Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., et al. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. *Archives Of General Psychiatry*, *62*(7), 799–805.

- Blackburn, R., Logan, C., Donnelly, J., & Renwick, S. (2008). Identifying psychopathic subtypes: Combining an empirical personality classification of offenders with the Psychopathy Checklist-Revised. *Journal of Personality Disorders, 22*(6), 604–622.
- Blair, J. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition, 57*(1), 1–29.
- Blair, J. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition, 14*(4), 698–718.
- Blair, J. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences, 11*(9), 387–392.
- Blair, J. (2008). Fine cuts of empathy and the amygdala: Dissociable deficits in psychopathy and autism. *The Quarterly Journal of Experimental Psychology, 61*(1), 157–170.
- Blair, J., Colledge, E., & Mitchell, D. (2001). Somatic markers and response reversal: Is there orbitofrontal cortex dysfunction in boys with psychopathic tendencies? *Journal of Abnormal Child Psychology, 29*(6), 499–511.
- Blair, J., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology, 34*(2), 192–198.
- Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., et al. (1996). Theory of mind in the psychopath. *Journal of Forensic Psychiatry & Psychology, 7*(1), 15–25.
- Blonigen, D., Hicks, B., Krueger, R., Patrick, C., & Iacono, W. (2005). Psychopathic personality traits: Heritability and genetic overlap with internalizing and externalizing psychopathology. *Psychological Medicine, 35*(05), 637–648.
- Bookstein, A., Klein, S., & Raita, T. (2001). Fuzzy hamming distance: A new

- dissimilarity measure. In A. Amir & G. Landau (Eds.), *Combinatorial pattern matching* (p. 86-97). Jerusalem, Israel.
- Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain*, *117*(6), 1241-1253. Available from <http://brain.oxfordjournals.org/content/117/6/1241.abstract>
- Boyd, R. (1988). Is the repeated prisoner's dilemma a good model of reciprocal altruism? *Ethology and Sociobiology*, *9*(2-4), 211-222.
- Boyd, R., & Lorberbaum, J. (1987). No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature*, *327*, 58-59.
- Brandt, J., Kennedy, W., Patrick, C., & Curtin, J. (1997). Assessment of psychopathy in a population of incarcerated adolescent offenders. *Psychological Assessment*, *9*, 429-435.
- Brinkley, C., Schmitt, W., Smith, S., & Newman, J. (2001). Construct validation of a self-report psychopathy scale: does Levenson's Self-report Psychopathy Scale measure the same constructs as Hare's psychopathy checklist-revised? *Personality and Individual Differences*, *31*(7), 1021-1038.
- Brosnan, S., & De Waal, F. (2003). Monkeys reject unequal pay. *Nature*, *425*(6955), 297-299.
- Brownell, H., Michel, D., Powelson, J., & Gardner, H. (1983). Surprise but not coherence: Sensitivity to verbal humor in right-hemisphere patients. *Brain and Language*, *18*(1), 20-27.
- Bryant, F., & Yarnold, P. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington, DC: American

- Psychological Association.
- Buller, D. (2009). Four fallacies of pop evolutionary psychology. *Scientific American Magazine*, 300(1), 74–81.
- Buss, D., & Duntley, J. (2003). An evolutionary psychological perspective. In R. Bloom & N. Dess (Eds.), *Evolutionary psychology and violence: A primer for policymakers and public policy advocates* (pp. 115–128). Westport, CT: Praeger Publishers.
- Byrne, R., & Whiten, A. (1992). Cognitive evolution in primates: Evidence from tactical deception. *Man*, 27(3), 609–627.
- Chagnon, N. (1988). Life histories, blood revenge, and warfare in a tribal population. *Science*, 239(4843), 985.
- Chiappori, P., Levitt, S., & Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, 92(4), 1138–1151.
- Child, D. (2006). *The essentials of factor analysis*. New York, NY: Continuum International Publishing Group.
- Clark, L., Manes, F., Antoun, N., Sahakian, B., & Robbins, T. (2003). The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia*, 41(11), 1474–1483.
- Cleckley. (1941). *The mask of sanity: An attempt to reinterpret the so-called psychopathic personality*. St Louis, MO: The CV Mosby Company.
- Coase, R. (1937). The nature of the firm. *Economica*, 386–405.
- Colman, A., & Wilson, J. (1997). Antisocial personality disorder: An evolutionary game theory analysis. *Legal and Criminological Psychology*, 2, 23–34.
- Cooke, D. (1998). Cross-cultural aspects of psychopathy. *Psychopathy: Antisocial, criminal, and violent behavior*, 260–276.

- Cooke, D., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological assessment, 13*(2), 171–188.
- Cooke, D., Michie, C., & Hart, S. (2006). Facets of clinical psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 415–436). New York, NY: The Guilford Press.
- Cox, J., & Swarthout, J. (n.d.). *EconPort: Creating and maintaining a knowledge commons*. Retrieved Econport website: <http://www.econport.org>. Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.
- Cunningham, M., & Reidy, T. (1998). Antisocial personality disorder and psychopathy: Diagnostic dilemmas in classifying patterns of antisocial behavior in sentencing evaluations. *Behavioral Sciences & the Law, 16*(3), 333–351.
- Cunningham, M., & Reidy, T. (2002). Violence risk assessment at federal capital sentencing. *Criminal Justice and Behavior, 29*(5), 512.
- Daly, M., & Wilson, M. (1988). *Homicide*. New York, NY: A. de Gruyter.
- Damasio, A. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York, NY: Putnam.
- Damasio, A. (2000). A neural basis for sociopathy. *Archives of General Psychiatry, 57*(2), 128–132.
- Dapretto, M., Davies, M., Pfeifer, J., Scott, A., Sigman, M., Bookheimer, S., et al. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience, 9*(1), 28–30.
- Darke, S., Kaye, S., & Finlay-Jones, R. (1998). Antisocial personality disorder, psychopathy and injecting heroin use. *Drug and Alcohol Dependence, 52*(1), 63–69.
- Dawkins, R. (1976). *The selfish gene*. (New York, NY: Oxford University Press,

- 2006)
- Deakin, J., Aitken, M., Robbins, T., & Sahakian, B. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society*, *10*(04), 590–598.
- Demski, J. (2003). Corporate conflicts of interest. *Journal of Economic Perspectives*, *17*(2), 51–72.
- Diamond, J. (2008). Annals of anthropology: Vengeance is ours retribution in New Guinea and beyond. *New Yorker Magazine*, 74.
- Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., , & Weingessel, A. (2009). e1071: Misc functions of the department of statistics (e1071) [Computer software manual]. (R package version 1.5-19)
- Dixit, A., & Nalebuff, B. (2008). *The art of strategy*. New York, NY: WW Norton & Co.
- Dugatkin, L. (1992). The evolution of the “con artist.”. *Ethology & Sociobiology*, *13*(1), 3–18.
- Dugatkin, L., & Wilson, D. (1991). Rover: A strategy for exploiting cooperators in a patchy environment. *American Naturalist*, *138*(3), 687–701.
- Edens, J., Marcus, D., Lilienfeld, S., & Poythress, N. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology*, *115*(1), 131.
- Ellis, L. (1988). Criminal behavior and r/K selection: An extension of gene-based evolutionary theory. *Personality and Individual Differences*, *9*(4), 697–708.
- Ellison, G. (1994). Cooperation in the prisoner’s dilemma with anonymous random matching. *The Review of Economic Studies*, *61*(3), 567–588.
- Erev, I., & Roth, A. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American*

- Economic Review*, 88(4), 848–881.
- Eysenck, H. (1987). The definition of personality disorders and the criteria appropriate for their description. *Journal of Personality Disorders*, 1(3), 211–219.
- Fecteau, S., Pascual-Leone, A., & Théoret, H. (2008). Psychopathy and the mirror neuron system: Preliminary findings from a non-psychiatric sample. *Psychiatry Research*, 160(2), 137.
- Forth, A., Kosson, D., & Hare, R. (2003). *Hare psychopathy checklist: Youth version (PCL: YV)*. North Tonawanda, NY: Multi-Health Systems.
- Fowles, D., & Dindo, L. (2006). A dual-deficit model of psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 14–34). New York, NY: The Guilford Press.
- Frank, R. (1996). What price the moral high ground? *Southern Economic Journal*, 63(1), 1–17.
- French, J. (1951). *The description of aptitude and achievement tests in terms of rotated factors*. Chicago University Press, Chicago.
- Frick, P. (1995). Callous-unemotional traits and conduct problems: A two-factor model of psychopathy in children. *Issues in Criminological & Legal Psychology*, 24, 47–51.
- Frick, P., & Hare, R. (2001). *The psychopathy screening device*. Toronto: Multi-Health Systems.
- Frick, P., Kimonis, E., Dandreaux, D., & Farell, J. (2003). The 4 year stability of psychopathic traits in non-referred youth. *Behavioral Sciences & the Law*, 21(6), 713–736.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2), 593.
- Gillstrom, B., & Hare, R. (1988). Language-related hand gestures in psychopaths.

- Journal of Personality Disorders*, 2(1), 21–27.
- Gintis, H. (2009). *Game theory evolving*. Princeton, NJ: Princeton University Press.
- Glenn, A., & Raine, A. (2009). The Immoral Brain. In J. Verplaetse, J. Jelle de Schrijver, S. Vanneste, & J. Braeckman (Eds.), *The moral brain: Essays on the evolutionary and neuroscientific aspects of morality* (pp. 45–68). Springer Verlag.
- Glicksohn, J., Naor-Ziv, R., & Leshem, R. (2007). Impulsive decision-making: Learning to gamble wisely? *Cognition*, 105(1), 195–205.
- Goethe, J. von. (1853). *Faust: a tragedy*. (translated by: Falck-Lebahn, K. London: Longman, Brown, & Green)
- Gordon, H., Baird, A., & End, A. (2004). Functional differences among those high and low on a trait measure of psychopathy. *Biological Psychiatry*, 56(7), 516–521.
- Gorsuch, R. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, 25(1), 33–39.
- Gray, J., & McNaughton, N. (2003). A theory of the behavioural inhibition system. *The Neuropsychology of Anxiety*, 1(9), 83–94.
- Grayson, B., & Stein, M. (1981). Attracting Assault: Victims. *Journal of Communication*, 31(1), 68–75.
- Greene, J., Sommerville, R., Nystrom, L., Darley, J., & Cohen, J. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105.
- Guay, J., Ruscio, J., Knight, R., & Hare, R. (2007). A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality. *Journal of Abnormal Psychology*, 116(4), 701.
- Guilford, J. (1956). The structure of intellect. *Psychological Bulletin*, 53(4), 267–293.
- Hamilton, W. (1964). The genetical evolution of social behaviour. *Journal of Theo-*

- retical Biology*, 7(1), 1–16.
- Hare, R. (1991). *The Hare Psychopathy Checklist*. Toronto: Multi-Health Systems.
- Hare, R. (1998). Psychopaths and their nature: Implications for the mental health and criminal justice systems. *Psychopathy: Antisocial, criminal, and violent behavior*, 188–212.
- Hare, R. (2003). *The Hare PCL-R*. Toronto: Multi-Health Systems.
- Hare, R., Harpur, T., Hakstian, A., Forth, A., Hart, S., & Newman, J. (1990). The revised psychopathy checklist: Reliability and factor structure. *Psychological Assessment*, 2(3), 338–341.
- Hare, R., & Jutai, J. (1988). Psychopathy and cerebral asymmetry in semantic processing. *Personality and Individual Differences*, 9(2), 329–337.
- Hare, R., & Neumann, C. (2006). The PCL-R assessment of psychopathy: development, structural properties, and new directions. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 58–88). New York, NY: The Guilford Press.
- Harpending, H., & Sobus, J. (1987). Sociopathy as an adaptation. *Ethology & Sociobiology*, 8(3), 63–72.
- Harpur, T., Hare, R., & Hakstian, A. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Psychological Assessment*, 1(1), 6–17.
- Harris, G., & Rice, M. (2006). Treatment of Psychopathy: a review of empirical findings. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 555–572). New York, NY: Guilford Press.
- Harris, G., Rice, M., Hilton, N., Lalumière, M., & Quinsey, V. (2007). Coercive and precocious sexuality as a fundamental aspect of psychopathy. *Journal of Personality Disorders*, 21(1), 1–27.
- Harris, G., Rice, M., & Quinsey, V. (1994). Psychopathy as a taxon: Evidence that

- psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology*, *62*(2), 387–397.
- Harris, G., Skilling, T., & Rice, M. (2001). The construct of psychopathy. *Crime & Justice*, *28*, 197.
- Heap, S., & Varoufakis, Y. (2004). *Game theory: A critical text*. London, UK: Routledge.
- Hein, G., & Singer, T. (2008). I feel how you feel but not always: The empathic brain and its modulation. *Current Opinion in Neurobiology*, *18*(2), 153–158.
- Hemphill, J., Hare, R., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology*, *3*(1), 139–170.
- Herpertz, S., & Koetting, K. (2005). Startle response in inpatients with borderline personality disorder vs. healthy controls. *Journal of Neural Transmission*, *112*(8), 1097–1106.
- Hirshleifer, J. (2001). On the emotions as guarantors of threats and promises. In J. Hirshleifer (Ed.), *The dark side of the force: economic foundations of conflict theory* (pp. 198–219). Cambridge, UK: Cambridge University Press.
- Hobbes, T., & Gaskin, J. (1651). *Leviathan*. (Edited by Gaskin, J.C.A. Oxford University Press: New York, NY. 1998)
- Honk, J. van, Hermans, E., Putman, P., Montagne, B., & Schutter, D. (2002). Defective somatic markers in sub-clinical psychopathy. *Neuroreport*, *13*(8), 1025.
- Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, *7*(12), 942–951.
- Intrator, J., Hare, R., Stritzke, P., Brichtswein, K., Dorfman, D., Harpur, T., et al. (1997). A brain imaging (single photon emission computerized tomography) study of semantic and affective processing in psychopaths. *Biological Psychiatry*,

- 42(2), 96–103.
- Ip, E. (2010). Interpretation of the three-parameter testlet response model and information function. *Applied Psychological Measurement, 34*(7), 467–482.
- Jablonka, E., & Lamb, M. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge, MA: The MIT Press.
- Jackson, P., Meltzoff, A., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage, 24*(3), 771–779.
- Johansson, S., Carlsson, B., & Boman, M. (n.d.). Modeling strategies as generous and greedy in prisoner's dilemma like games. In M. B., X. Yao, N. C., J. Kim, & T. Furuhashi (Eds.), *Simulated evolution and learning*.
- Justus, A., & Finn, P. (2007). Startle modulation in non-incarcerated men and women with psychopathic traits. *Personality and Individual Differences, 43*(8), 2057–2071.
- Kaiser, J. (2009). Private money, public disclosure. *Science, 325*(5936), 28.
- Karpman, B. (1948). The myth of the psychopathic personality. *American Journal of Psychiatry, 104*(9), 523.
- Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences, 11*(5), 194–196.
- Keysers, C., & Gazzola, V. (2009). Unifying social cognition. In J. Pineda (Ed.), *Mirror neuron systems: The role of mirroring processes in social cognition* (pp. 3–28). New York, NY: Humana Press/Springer Science.
- Kiehl, K., Hare, R., McDonald, J., & Brink, J. (1999). Semantic and affective processing in psychopaths: An event-related potential (ERP) study. *Psychophysiology, 36*(06), 765–774.

- Kiehl, K., Smith, A., Hare, R., Mendrek, A., Forster, B., Brink, J., et al. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry*, *50*(9), 677–684.
- Kiehl, K., Smith, A., Mendrek, A., Forster, B., Hare, R., & Liddle, P. (2004). Temporal lobe abnormalities in semantic processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Psychiatry Research: Neuroimaging*, *130*(3), 297–312.
- Kinner, S. (2003). Psychopathy as an adaptation implications for society and social policy. In R. Bloom & N. Dess (Eds.), *Evolutionary psychology and violence: A primer for policymakers and public policy advocates* (pp. 57–82). Westport, CT: Praeger Publishers.
- Kitcher, P. (1985). *Vaulting ambition*. Cambridge, MA.: MIT Press.
- Kiyonari, T., Tanida, S., & Yamagishi, T. (2000). Social exchange and reciprocity: Confusion or a heuristic? *Evolution and Human Behavior*, *21*(6), 411–427.
- Kosson, D., Suchy, Y., Mayer, A., & Libby, J. (2002). Facial affect recognition in criminal psychopaths. *Emotion*, *2*(4), 398–411.
- Larsson, H., Andershed, H., & Lichtenstein, P. (2006). A genetic factor explains most of the variation in the psychopathic personality. *Journal of Abnormal Psychology*, *115*(2), 221.
- Lawrence, A., Clark, L., Labuzetta, J., Sahakian, B., & Vyakarnum, S. (2008). The innovative brain. *Nature*, *456*(7219), 168–169.
- Lawrence, A., Luty, J., Bogdan, N., Sahakian, B., & Clark, L. (2009). Problem gamblers share deficits in impulsive decision-making with alcohol-dependent individuals. *Addiction*, *104*(6), 1006–1015.
- Lebiere, C., & West, R. (1999). A dynamic ACT-R model of simple games. In

- Proceedings of the twenty-first annual conference of the cognitive science society* (pp. 296–301). Mahwah, NJ: Lawrence Erlbaum.
- Levenson, M. (1990). Risk taking and personality. *Journal of Personality and Social Psychology, 58*(6), 1073–1080.
- Levenson, M., Kiehl, K., & Fitzpatrick, C. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology, 68*, 151–151.
- Lilienfeld, S. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review, 14*(1), 17–38.
- Lilienfeld, S., & Widows, M. (2005). *Professional manual for the Psychopathic Personality Inventory-Revised (PPI-RR)*. Lutz, FL: Psychological Assessment Resources.
- Lorberbaum, J. (1994). No strategy is evolutionarily stable in the repeated prisoner's dilemma. *Journal of Theoretical Biology, 168*(2), 117–130.
- Lorenz, A., & Newman, J. (2002). Do emotion and information processing deficiencies found in Caucasian psychopaths generalize to African-American psychopaths? *Personality and Individual Differences, 32*(6), 1077–1086.
- Lykken, D. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology, 55*(1), 6–10.
- Lynam, D., Whiteside, S., & Jones, S. (1999). Self-reported psychopathy: A validation study. *Journal of Personality Assessment, 73*(1), 110–132.
- Machiavelli, N. (1532). *The Prince*. (Edited by McMahon, R. and Marriott, WK. New York, NY: Sterling Publishing.2008)
- MacLean, P. (1990). *The triune brain in evolution: Role in paleocerebral functions*. New York, NY: Plenum Press.
- Mailath, G. (1998). Do people play Nash equilibrium? Lessons from evolutionary

- game theory. *Journal of Economic Literature*, 36(3), 1347–1374.
- Maindonald, J., & Braun, W. J. (2009). Daag: Data analysis and graphics data and functions [Computer software manual]. Available from <http://CRAN.R-project.org/package=DAAG> (R package version 1.01)
- Marcus, D., John, S., & Edens, J. (2004). A taxometric analysis of psychopathic personality. *Journal of Abnormal Psychology*, 113, 626–635.
- Maynard-Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.
- McCotter, T. (2010). Hitting streaks dont obey your rules. *Chance*, 23(4), 52–57.
- McCulloch, C. (2010). freqmap: Frequency moving average plots (map) of multinomial data by a continuous covariate [Computer software manual]. Available from <http://CRAN.R-project.org/package=freqMAP> (R package version 0.2)
- McDermott, P., Alterman, A., Cacciola, J., Rutherford, M., Newman, J., & Mulholland, E. (2000). Generality of Psychopathy Checklist-Revised factors over prisoners and substance-dependent patients. *Journal of Consulting and Clinical Psychology*, 68(1), 181–186.
- Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavior and Brain Sciences*, 18–523.
- Meehl, P. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, 60(1), 117–174.
- Meehl, P. (1995). Bootstraps taxometrics. Solving the classification problem in psychopathology. *The American Psychologist*, 50(4), 266.
- Meehl, P. (2004). What's in a taxon? *Journal of Abnormal Psychology*, 113(1), 39–43.
- Mitchell, D., Colledge, E., Leonard, A., & Blair, J. (2002). Risky decisions and response reversal: is there evidence of orbitofrontal cortex dysfunction in psy-

- chopathic individuals? *Neuropsychologia*, *40*(12), 2013–2022.
- Mokros, A., Menner, B., Eisenbarth, H., Alpers, G., Lange, K., & Osterheider, M. (2008). Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *Journal of Abnormal Psychology*, *117*(2), 406.
- Montañes, R., Lucas, T. de, & Martín, R. (2003). Antisocial personality disorder evaluation with the prisoner's dilemma. *Actas españolas de psiquiatría*, *31*(6), 307.
- Murrie, D., Marcus, D., Douglas, K., Lee, Z., Salekin, R., & Vincent, G. (2007). Youth with psychopathy features are not a discrete class: A taxometric analysis. *Journal of Child Psychology and Psychiatry*, *48*(7), 714–723.
- Murzynski, J., & Degelman, D. (1996). Body language of women and judgments of vulnerability to sexual assault. *Journal of Applied Social Psychology*, *26*, 1617–1626.
- Nesse, R. (1990). Evolutionary explanations of emotions. *Human Nature*, *1*(3), 261–289.
- Neumann, C., & Hare, R. (2008). Psychopathic traits in a large community sample: Links to violence, alcohol use, and intelligence. *Journal of Consulting and Clinical Psychology*, *76*(5), 893–899.
- Neumann, C., Kosson, D., Forth, A., et al. (2006). Factor structure of the Hare Psychopathy Checklist: Youth Version (PCL: YV) in incarcerated adolescents. *Psychological Assessment*, *18*(2), 142–154.
- Neumann, C., Kosson, D., & Salekin, R. (2007). Exploratory and confirmatory factor analysis of the psychopathy construct: Methodological and conceptual issues. *The psychopath: Theory, research, and practice*, 79–104.
- Newman, J. (1998). Psychopathic behavior: An information processing perspective. In D. Cooke, A. Forth, & R. Hare (Eds.), *Psychopathy: Theory, research and*

- implications for society* (pp. 81–104). New York, NY: Springer.
- Newman, J., & Kosson, D. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology, 95*(3), 252–256.
- Newman, J., Patterson, C., & Kosson, D. (1987). Response perseveration in psychopaths. *Journal of Abnormal Psychology, 96*(2), 145–148.
- Nowak, M. A. (2006). *Evolutionary dynamics : Exploring the equations of life*. Cambridge, MA: Belknap Press.
- Ooms, J. C. (2009). yeroon.net/lme4: A web interface for the r package lme4 [Computer software manual]. Available from <http://www.yeroon.net/lme4> (Version 0.1)
- Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. New York, NY: Oxford University Press.
- Panksepp, J., & Panksepp, J. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition, 6*(2), 108–131.
- Patrick, C. (1994). Emotion and psychopathy: Startling new insights. *Psychophysiology, 31*(4), 319–330.
- Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*(1), 176–180.
- Pilcher, C., Wong, J., & Pillai, S. (2008). Inferring HIV transmission dynamics from phylogenetic sequence relationships. *PLoS Med, 5*(3), e69. (Retrieved from: <http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0050069>)
- Pinel, P. (1818). *Nosographie philosophique, ou, La méthode de l'analyse appliquée a la médecine*. Paris: chez JA Brosson.
- Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-Plus*. New York, NY: Springer.

- Pinker, S. (2000). *The language instinct: How the mind creates language*. Harper Perennial Modern Classics.
- Poundstone, W. (1993). *Prisoner's dilemma*. New York, NY: Doubleday.
- R Development Core Team. (2009). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from <http://www.R-project.org> (ISBN 3-900051-07-0)
- Raine, A., Lencz, T., Bihrl, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, *57*(2), 119.
- Relman, A. (2001). Separating continuing medical education from pharmaceutical marketing. *Journal of the American Medical Association*, *285*(15), 2009.
- Ricoeur, P. (1986). *Fallible man*. New York, NY: Fordham University Press.
- Rilling, J., Glenn, A., Jairam, M., Pagnoni, G., Goldsmith, D., Elfenbein, H., et al. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological Psychiatry*, *61*(11), 1260–1271.
- Rizzolatti, G., & Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194.
- Rush, B. (1972). *Two essays on the mind*. New York, NY: Brunner/Mazel.
- Sakaguchi, K., & Hasegawa, T. (2006). Person perception through gait information and target choice for sexual advances: Comparison of likely targets in experiments and real life. *Journal of Nonverbal Behavior*, *30*(2), 63–85.
- Schmitt, W., Brinkley, C., & Newman, J. (1999). Testing Damasio's somatic marker hypothesis with psychopathic individuals: Risk takers or risk averse? *Journal of Abnormal Psychology*, *108*(3), 538.
- Schmitt, W., & Newman, J. (1999). Are all psychopathic individuals low-anxious? *Journal of Abnormal Psychology*, *108*, 353–358.

- Seto, M., Khattar, N., Lalumière, M., & Quinsey, V. (1997). Deception and sexual strategy in psychopathy. *Personality and Individual Differences, 22*(3), 301–307.
- Seto, M., Lalumière, M., & Quinsey, V. (1995). Sensation seeking and males' sexual strategy. *Personality and Individual Differences, 19*(5), 669–675.
- Shubin, N., Daeschler, E., & Jenkins, F. (2006). The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature, 440*(7085), 764–771.
- Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Skilling, T., Harris, G., Rice, M., & Quinsey, V. (2002). Identifying persistently antisocial offenders using the Hare Psychopathy Checklist and DSM antisocial personality disorder criteria. *Psychological Assessment, 14*(1), 27–38.
- Skilling, T., Quinsey, V., & Craig, W. (2001). Evidence of a taxon underlying serious antisocial behavior in boys. *Criminal Justice and Behavior, 28*(4), 450.
- Smith, D. L. (2004). *Why we lie: The evolutionary roots of deception and the unconscious mind* (1st ed.). New York, NY: St. Martin's Press.
- Smith, D. L. (2009). *The most dangerous animal: Human nature and the origins of war*. New York, NY.
- Stearns, S. C., & Koella, J. C. (2008). *Evolution in health and disease* (2nd ed.). New York, NY: Oxford University Press.
- Taylor, J., Loney, B., Bobadilla, L., Iacono, W., & McGue, M. (2003). Genetic and environmental influences on psychopathy trait dimensions in a community sample of male twins. *Journal of Abnormal Child Psychology, 31*(6), 633–645.
- Tenbrunsel, A., & Messick, D. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research, 17*(2), 223–236.
- Tinbergen, N. (2005). On aims and methods of ethology. *Animal Biology—Leiden,*

- 55(4), 297–322.
- Tooby, J., & Cosmides, L. (n.d.). Mapping the evolved functional organization of mind and brain. In M. Gazzaniga (Ed.), *The new cognitive neurosciences*.
- Tooby, J., & Cosmides, L. (2009). Conceptual foundations of evolutionary psychology. In A. Rosenberg & R. Arp (Eds.), *Philosophy of biology: An anthology* (pp. 375–386). Oxford, UK: Blackwell.
- Tooke, W., & Camire, L. (1991). Patterns of deception in intersexual and intrasexual mating strategies. *Ethology and Sociobiology*, 12(5), 345–364.
- Trivers, R. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35–57.
- Trivers, R. (2002). *Natural selection and social theory: Selected papers of Robert Trivers*. New York, NY: Oxford University Press.
- Vanman, E., Mejia, V., Dawson, M., Schell, A., & Raine, A. (2003). Modification of the startle reflex in a community sample: Do one or two dimensions of psychopathy underlie emotional processing? *Personality and Individual Differences*, 35(8), 2007–2021.
- Veit, R., Flor, H., Erb, M., Hermann, C., Lotze, M., Grodd, W., et al. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neuroscience Letters*, 328(3), 233–236.
- Vitacco, M., Neumann, C., & Jackson, R. (2005). Testing a four-factor model of psychopathy and its association with ethnicity, gender, intelligence, and violence. *Journal of Consulting and Clinical Psychology*, 73, 466–476.
- Wade, N. (2010). *Harvard researcher may have fabricated data*. Retrieved August 27, 2010, from NYTimes website: http://www.nytimes.com/2010/08/28/science/28harvard.html?_r=1. New York Times.
- Walker, N., & McCabe, S. (1973). *Crime and insanity in England: New Solutions*

- and new problems* (Vol. 2). Edinburgh, UK: Edinburgh University Press.
- Walters, G., Gray, N., Jackson, R., Sewell, K., Rogers, R., Taylor, J., et al. (2007). A taxometric analysis of the Psychopathy Checklist: Screening Version (PCL:SV): Further evidence of dimensionality. *Psychological Assessment, 19*(3), 330–339.
- West, R., Lebiere, C., & Bothell, D. (2006). Cognitive architectures, game playing, and human evolution. In R. Sun (Ed.), *Cognition and multi-agent interaction: From cognitive modeling to social simulation* (pp. 103–123). New York, NY: Cambridge University Press.
- West, R., Stewart, T., Lebiere, C., & Chandrasekharan, S. (n.d.). Stochastic resonance in human cognition: ACT-R versus game theory, associative neural networks, recursive neural networks, Q-learning, and humans. In B. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *27th annual meeting of the cognitive science society*.
- Wheeler, S., Book, A., & Costello, K. (2009). Psychopathic traits and perceptions of victim vulnerability. *Criminal Justice and Behavior, 36*(6), 635.
- Widiger, T., Cadoret, R., Hare, R., Robins, L., Rutherford, M., Zanarini, M., et al. (1996). DSM-IV antisocial personality disorder field trial. *Journal of Abnormal Psychology, 105*, 3–16.
- Widiger, T., & Lynam, D. (1998). Psychopathy and the five-factor model of personality. In T. Millon, E. Simonsen, M. Birket-Smith, & R. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behavior* (pp. 171–187). New York, NY: Guilford Press.
- Widom, C. (1976). Interpersonal conflict and cooperation in psychopaths. *Journal of Abnormal Psychology, 85*(3), 330–334.
- Wilkinson, G. (1984). Reciprocal food sharing in the vampire bat. *Nature, 308*(5955),

181–184.

- Williams, J., Oehlert, G., Carlis, J., & Pusey, A. (2004). Why do male chimpanzees defend a group range? *Animal behaviour*, *68*(3), 523–532.
- Williamson, D., Womble, L., Smeets, M., Netemeyer, R., Thaw, J., Kutlesic, V., et al. (2002). Latent structure of eating disorder symptoms: A factor analytic and taxometric investigation. *American Journal of Psychiatry*, *159*(3), 412.
- Williamson, S., Harpur, T., & Hare, R. (1991). Abnormal processing of affective words by psychopaths. *Psychophysiology*, *28*(3), 260–273.
- Wilson, E. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Belknap.
- Wrangham, R. (1999). Evolution of coalitionary killing. *Yearbook of Physical Anthropology*, *42*, 1–30.
- Wrangham, R., & Peterson, D. (1996). *Demonic males: Apes and the origins of human violence*. Boston, MA: Houghton Mifflin.

Appendices

Appendix A

The Levenson Self-Report Psychopathy Scale

1. Success is based on survival of the fittest; I am not concerned about the losers.
2. For me, what's right is whatever I can get away with.
3. In today's world, I feel justified in doing anything I can get away with to succeed.
4. My main purpose in life is getting as many goodies as I can.
5. Making a lot of money is my most important goal.
6. I let others worry about higher values; my main concern is with the bottom line.
7. People who are stupid enough to get ripped off usually deserve it.
8. Looking out for myself is my top priority.
9. I tell other people what they want to hear so that they will do what I want them to do.
10. I would be upset if my success came at someone else's expense.
11. I often admire a really clever scam.
12. I make a point of trying not to hurt others in pursuit of my goals.
13. I enjoy manipulating other people's feelings.
14. I feel bad if my words or actions cause someone else to feel emotional pain.
15. Even if I were trying very hard to sell something, I wouldn't lie about it.

16. Cheating is not justified because it is unfair to others.
17. I find myself in the same kinds of trouble, time after time.
18. I am often bored.
19. I find that I am able to pursue one goal for a long time.
20. I don't plan anything very far in advance.
21. I quickly lose interest in tasks I start.
22. Most of my problems are due to the fact that other people just don't understand me.
23. Before I do anything, I carefully consider the possible consequences.
24. I have been in a lot of shouting matches with other people.
25. When I get frustrated, I often "let off steam" by blowing my top.
26. Love is overrated.

Appendix B

Physical Set-up

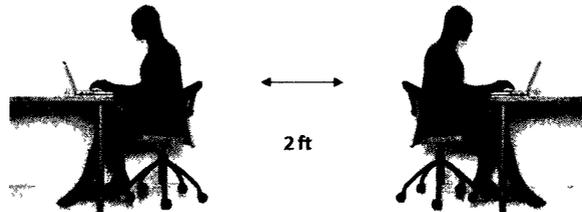


Figure B.1: Positioning of Participants

Appendix C

Prisoner Dilemma Instructions

In Game #1 you will be playing a computer game against each other (the other student in the pair) for 40 rounds. The game is played on a computer terminal in which you are paired with another student and your task is to select a move “S” or “C”. Take a look at your monitor and find the message that tells you whether you are the row or column player.

Now hover your mouse over a column (or row) and notice that your payoff appears in bold font. Try moving your mouse to the other choice and see that your payoff appears in bold too. Notice that your payoff differs for each move and also differs depending on the move of the other player, which you do not know until the end of a round.

To select a move, click on the a row (or column). Once you click on a choice, you cannot change it so make sure you choose well. After you and your partner have both moved, you will get feedback about the amount you won. Take note especially of the figure on the upper left hand corner of the window which shows your accumulated winnings so far. Your winnings at the end of the 40 rounds will be converted to its equivalent value in cents (1 point = 1 cent).

I will now test to your understanding of the logic of the game.

1. Suppose the Row Player in [in Figure 3] plays “S” and the Column Player, plays “C”, how much does the Row Player win? 0. The Column Player? 5
2. In a pair of numbers like 2.5, 0, how much is the column player’s payoff? 0
3. Are you allowed to communicate your intended choice to the other? No.

We will now play 5 test-rounds to familiarize you with the game.

Appendix D

Iowa Gambling Task Instructions

In front of you on the screen, there are 4 decks of cards A, B, C, and D. I want you to select one card at a time, by clicking on the card, from any deck you choose. Each time you select a card, the computer will tell you that you won some money. I don't know how much money you will win. You will find out as we go along. Every time you win, the green bar gets bigger. Every so often, however, when you click on a card, the computer tells you that you won some money, but then it says that you lost some money too. I don't know when you will lose, or how much you will lose. You will find out as we go along. Every time you lose, the green bar gets smaller. You are absolutely free to switch from one deck to the other at any time, as often as you wish. The goal of the game is to win as much money as possible, and if you can't win, avoid losing money as much as possible. You won't know when the game will end. You must keep on playing until the computer stops. I am going to give you this \$2000 credit, the green bar, to start the game. The red bar here is a reminder of how much money you borrowed to play the game, and how much money you have to pay back before we can see how much you won or lost. It is important to know that just like in a real card game, the computer does not change the order of the cards after the game starts. You may not be able to figure out exactly when you will lose money, but the game is fair. The computer does not make you lose money at random, or

make you lose money based on the last card you picked. Also, each deck contains an equal number of cards of each color, so the color of the cards does not tell you which decks are better in this game. So you must not try to figure out what the computer is doing. All I can say is that some decks are works than the others. You may find all of them bad, but some are worse than the others. No matter how much you find yourself losing, you can still win if you stay away from the worst decks. Please treat the play money in this game as real money, and any decision on what to do with it should be made as if you were using your own money.

Appendix E

Distribution of 'Cooperate' Moves

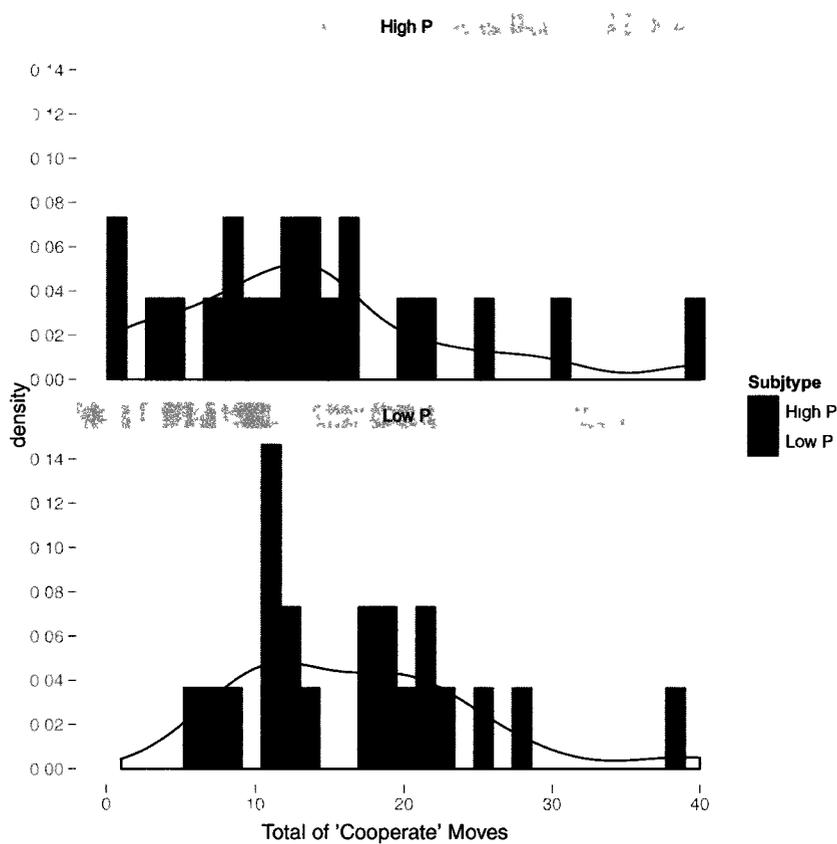


Figure E.1: 'Cooperate' Moves by *Subject Type*

Appendix F

Distribution of PD Outcomes

Table F.1: Self-Advantageous Defection in the Prisoner Dilemma

Match #	Player Type	Player	CC	CD	DC	DD	Self-Advantageous Defection Index
1	Low P	1	4	8	13	15	1 62
1	Low P	2	4	13	8	15	0 61
2	Low P	1	10	12	11	7	0 91
2	Low P	2	10	11	12	7	1 09
3	Low P	1	5	20	2	13	0 1
3	Low P	2	5	2	20	13	10
4	High P	1	22	0 1	4	14	40
4	High P	2	22	4	0 1	14	025
5	High P	1	0	1	1	38	1
5	High P	2	0	1	1	38	1
6	Low P	1	6	5	13	16	2 6
6	Low P	2	6	13	5	16	0 38
7	Low P	1	1	5	11	23	2 2
7	Low P	2	1	11	5	23	0 45
8	High P	1	2	8	12	18	1 5
8	High P	2	2	12	8	18	0 66
9	High P	1	7	9	7	17	0 77
9	High P	2	7	7	9	17	1 28
10	Low P	1	4	14	9	13	0 64
10	High P	2	4	9	14	13	1 55
11	Low P	1	23	5	7	5	1 4
11	High P	2	23	7	5	5	0 71
12	Low P	1	4	10	12	14	1 2
12	High P	2	4	12	10	14	0 83
13	Low P	1	3	16	6	15	0 37
13	Low P	2	3	6	16	15	2 66
14	Low P	1	2	9	7	22	0 77
14	High P	2	2	7	9	22	1 28
15	Low P	1	2	9	9	20	1
15	Low P	2	2	9	9	20	1
16	Low P	1	0	20	3	17	0 15
16	High P	2	0	3	20	17	6 66
17	Low P	1	39	0 1	1	0	10
17	High P	2	39	1	0 1	0	0 1
18	Low P	1	9	14	11	6	0 78
18	High P	2	9	11	14	6	1 27
19	High P	1	2	9	6	23	0 66
19	High P	2	2	6	9	23	1 5
20	High P	1	5	10	7	18	0 7
20	High P	2	5	7	10	18	1 42
21	High P	1	0	7	5	28	0 71
21	High P	2	0	5	7	28	1 4

Appendix G

Low P Streaks of Bad Deck Choices

Appendix H

Equilibrium frequencies of High P, Low P

The payoff matrix for pairwise matches is given by:

$$\begin{array}{cc} & \begin{array}{cc} \text{Low P} & \text{High P} \end{array} \\ \begin{array}{c} \text{Low P} \\ \text{High P} \end{array} & \left(\begin{array}{cc} 53.5 & 48 \\ 68 & 47.5 \end{array} \right) \end{array}$$

We can simplify the calculation by subtracting the diagonals (Nowak, 2006) (p. 58) from their off-diagonal entries yielding an equivalent matrix:

$$\begin{array}{cc} & \begin{array}{cc} \text{Low P} & \text{High P} \end{array} \\ \begin{array}{c} \text{Low P} \\ \text{High P} \end{array} & \left(\begin{array}{cc} 0 & 0.5 \\ 14.5 & 0 \end{array} \right) \end{array}$$

Let p be the frequency of Low P strategists and $1-p$ be the High P strategists. We can find the equilibrium points of 6.1 when the payoffs of Low P strategies equal High P strategists—i.e. each type's payoff equals the average payoff ϕ . So our equation is:

$$0 \times p + 0.5 \times (1 - p) = 14.5 \times p + 0 \times (1 - p) \quad (\text{H.1})$$

The equilibrium frequencies are: $p = .03$ and $1 - p = .97$, in perfect agreement with the simulation in Matlab.

Appendix I

Equilibrium frequencies of High P, Low P, with varying Initial Conditions

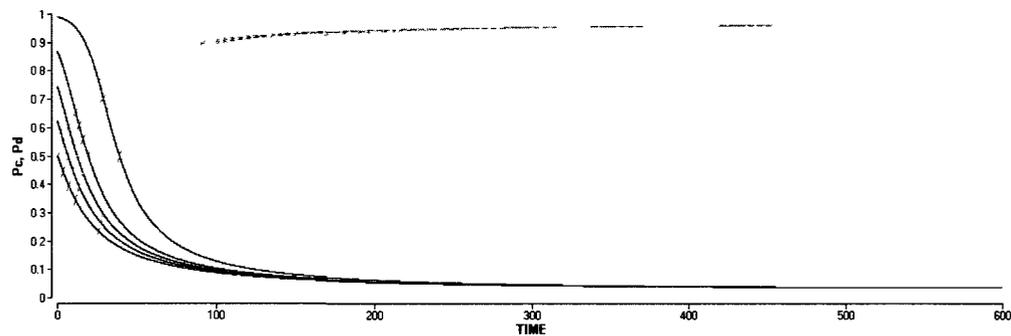


Figure I.1: Different Starting Values of High P, Low P Strategies

Appendix J

Hamming Distances

		Strategy																				HD	
Strategy		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>	<i>m</i>	<i>n</i>	<i>o</i>	<i>p</i>	<i>q</i>	<i>r</i>	<i>s</i>	<i>t</i>	<i>u</i>	
	<i>a</i>	0	23	24	28	22	21	22	17	21	29	4	23	28	26	21	28	23	18	24	24	23	449
	<i>b</i>	23	0	9	15	17	10	19	12	14	8	25	2	15	13	12	29	4	39	9	13	6	294
	<i>c</i>	24	9	0	16	18	13	12	17	19	13	24	11	20	12	13	24	13	30	12	16	11	327
	<i>d</i>	28	15	16	0	20	21	20	19	21	11	26	15	16	16	17	18	17	24	16	18	21	375
	<i>e</i>	22	17	18	20	0	17	22	25	15	19	20	15	16	10	21	20	19	24	12	16	15	363
	<i>f</i>	21	10	13	21	17	0	21	18	18	16	23	10	17	15	14	25	10	31	13	21	12	346
	<i>g</i>	22	19	12	20	22	21	0	17	21	19	20	19	22	18	17	20	19	20	22	20	19	389
	<i>h</i>	17	12	17	19	25	18	17	0	16	14	21	12	23	23	16	27	14	29	15	15	14	364
	<i>i</i>	21	14	19	21	15	18	21	16	0	18	21	14	19	15	18	25	18	25	17	17	16	368
	<i>j</i>	29	8	13	11	19	16	19	14	18	0	31	6	15	17	16	23	8	33	13	15	12	336
	<i>k</i>	4	25	24	26	20	23	20	21	21	31	0	25	28	22	23	24	27	14	24	26	25	453
	<i>l</i>	23	2	11	15	15	10	19	12	14	6	25	0	13	13	14	29	4	39	9	13	6	292
	<i>m</i>	28	15	20	16	16	17	22	23	19	15	28	13	0	16	19	18	13	26	18	14	13	369
	<i>n</i>	26	13	12	16	10	15	18	23	15	17	22	13	16	0	15	20	17	26	10	16	15	335
	<i>o</i>	21	12	13	17	21	14	17	16	18	16	23	14	19	15	0	23	16	27	21	17	14	354
	<i>p</i>	28	29	24	18	20	25	20	27	25	23	24	29	18	20	23	0	29	10	24	20	25	461
	<i>q</i>	23	4	13	17	19	10	19	14	18	8	27	4	13	17	16	29	0	37	11	15	8	322
	<i>r</i>	18	39	30	24	24	31	20	29	25	33	14	39	26	26	27	10	37	0	32	28	35	547
	<i>s</i>	24	9	12	16	12	13	22	15	17	13	24	9	18	10	21	24	11	32	0	14	9	325
	<i>t</i>	24	13	16	18	16	21	20	15	17	15	26	13	14	16	17	20	15	28	14	0	9	347
	<i>u</i>	23	6	11	21	15	12	19	14	16	12	25	6	13	15	14	25	8	35	9	9	0	308

Table J.1: Hamming distances of High P strategies

		Strategy																				HD
Strategy		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	0	20	21	15	14	20	16	21	17	20	27	21	23	13	17	12	16	13	19	19	23	367
b	20	0	19	17	18	24	20	21	17	16	19	23	23	19	21	22	22	19	25	17	17	399
c	21	19	0	18	25	21	19	20	20	21	14	16	18	22	18	19	21	22	20	22	24	400
d	15	17	18	0	15	17	23	20	14	23	28	20	18	16	18	13	17	14	20	16	24	366
e	14	18	25	15	0	22	28	19	15	16	35	17	23	11	19	16	16	13	15	13	21	371
f	20	24	21	17	22	0	24	17	15	30	21	17	23	21	19	12	20	19	19	21	21	403
g	16	20	19	23	28	24	0	21	23	18	11	21	19	25	15	22	16	25	23	25	13	407
h	21	21	20	20	19	17	21	0	18	23	20	20	24	16	16	23	19	22	16	26	20	402
i	17	17	20	14	15	15	23	18	0	25	28	16	20	14	18	15	13	16	20	18	22	364
j	20	16	21	23	16	30	18	23	25	0	21	23	19	19	17	22	20	19	15	13	17	397
k	27	19	14	28	35	21	11	20	28	21	0	22	18	32	20	29	25	30	28	30	18	476
l	21	23	16	20	17	17	21	20	16	23	22	0	16	22	18	17	15	16	18	22	22	382
m	23	23	18	18	23	23	19	24	20	19	18	16	0	18	22	21	17	16	20	18	24	400
n	13	19	22	16	11	21	25	16	14	19	32	22	18	0	22	19	15	10	14	16	20	364
o	17	21	18	18	19	19	15	16	18	17	20	18	22	22	0	19	19	20	18	20	16	372
p	12	22	19	13	16	12	22	23	15	22	29	17	21	19	19	0	18	13	19	15	21	367
q	16	22	21	17	16	20	16	19	13	20	25	15	17	15	19	18	0	19	17	17	19	361
r	13	19	22	14	13	19	25	22	16	19	30	16	16	10	20	13	19	0	16	16	22	360
s	19	25	20	20	15	19	23	16	20	15	28	18	20	14	18	19	17	16	0	18	24	384
t	19	17	22	16	13	21	25	26	18	13	30	22	18	16	20	15	17	16	18	0	20	382
u	23	17	24	24	21	21	13	20	22	17	18	22	24	20	16	21	19	22	24	20	0	408

Table J.2: Hamming distances of Low P strategies

Appendix K

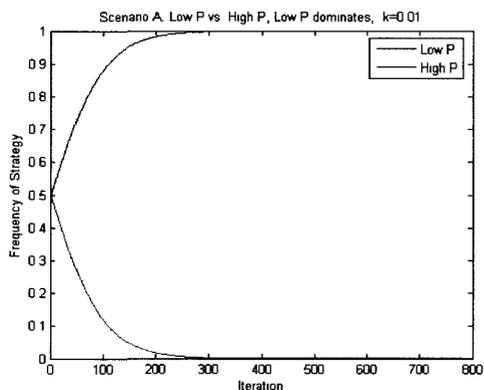
R Code for the CGT LMM

```
fm1<-lmer(Staketrans~ 1+BoxRatio+(BoxRatio|Treat) , data=Z)
fm2<-lmer(Staketrans~ 1+BoxRatio+(BoxRatio-1|LevBinned:Treat),data=Z)
fm3<-lmer(Staketrans~ 1+BoxRatio+(BoxRatio|LevBinned:Treat),data=Z)
fm4<-lmer(Staketrans~ 1+BoxRatio+(1|LevBinned:Treat),data=Z)
fm5<-lmer(Staketrans~1+BoxRatio+risktaking+(BoxRatio|LevBinned:Treat),data=Z)
fm8<-lmer(Staketrans~1+BoxRatio+risktaking+riskadj+
  (BoxRatio|LevBinned:Treat), data=Z))
```

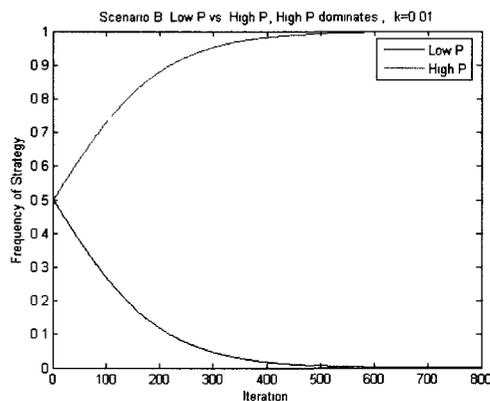
Appendix L

Evolutionary Scenarios in 2×2 matrix games

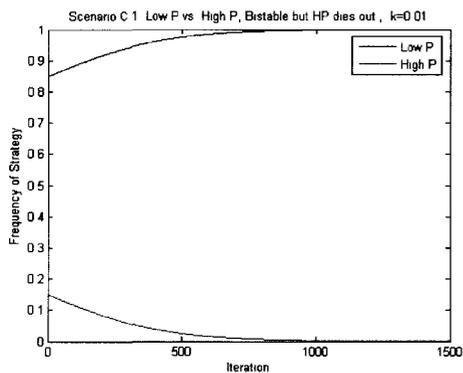
<i>Payoffs</i>	Low P	High P
Low P	a	b
High P	c	d



Low P dominates High P: if $a > c$ AND $b > d$



High P dominates Low P: if $a < c$ AND $b < d$



Low P and High P are bistable: if $a > c$ AND $b < d$

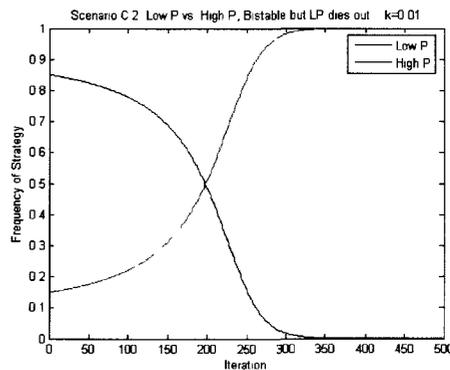
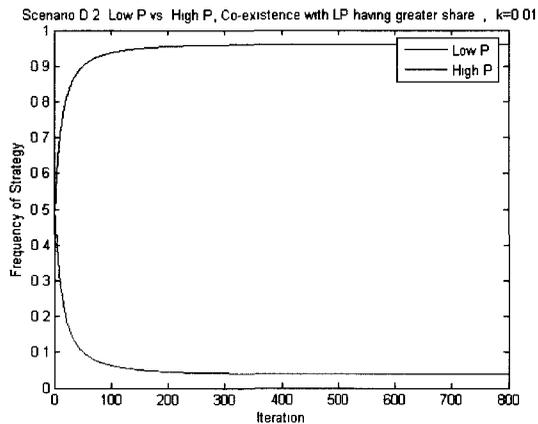
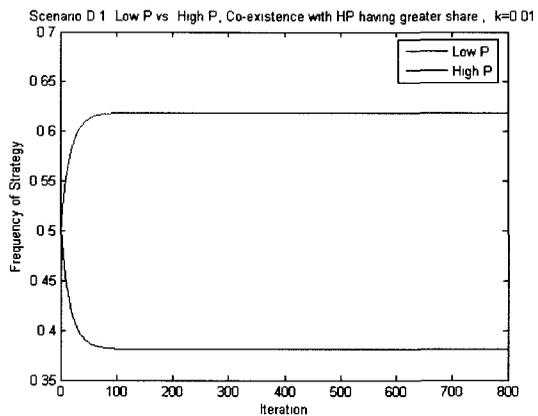
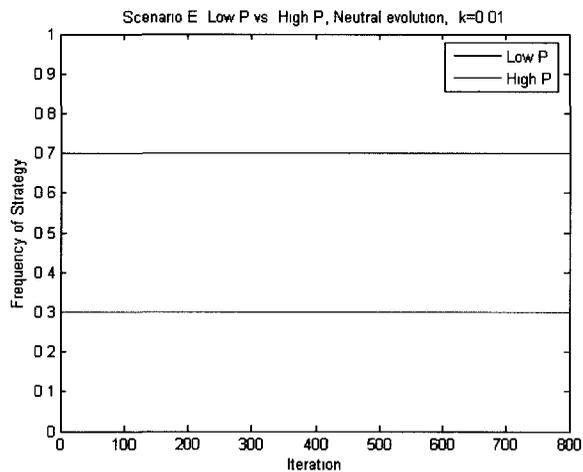


Figure L.1: Scenarios 1 to 4



Low P and High P co-exist: if $a < c$ AND $b > d$



Low P and High P are neutral: $a = c$ AND $b = d$

Figure L.2: Scenarios 5 to 7