THE UNIVERSITY OF CHICAGO

THE NEUROSCIENCE OF JUSTICE SENSITIVITY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

KEITH JONATHON YODER

CHICAGO, ILLINOIS

JUNE 2017

This work is dedicated to my parents, Darrel and Candy Yoder, for teaching me to seek truth with dedication and commitment, and to my wife, Laura, for her unwavering commitment to justice.

Table of Contents

List of Tables viii

List of Figures x

Acknowledgements xii

Abstract xiii

Chapter 1: Introduction 1

1.1 Justice as a Universal Motive 4

1.2 The Economics of Justice 6

1.3 Justice as a Personality Trait 11

Chapter 2: Behavioral Effects of Justice Sensitivity and Perspective 17

2.1 Introduction 17

2.2 Methods 21

2.2.1 Participants 21

2.2.2. Dispositional Measures 22

2.2.3. Stimuli and Tasks 22

2.2.4. Data Analysis 26

2.3 Results and Discussion 28

2.4 Appendix A: Chapter Two Tables and Figures 34

Chapter 3: Dispositional Influences on Fairness Judgments and Behaviors 49

3.1 Introduction 49

3.2. General Methods 49

3.1.1 Survey administration and participant recruitment 49

3.3. Between-Group Effects of Justice Perspectives on Fairness Judgments 50

3.3.1 Introduction 50
3.3.2 Methods 51
3.3.3 Results and Discussion 56
3.4 Within-Participant Effects of Justice Perspectives on Fairness Judgments 58
3.4.1 Introduction 58
3.4.2 Methods 58
3.4.3 Results and Discussion 61
3.5 Resolving Effects of Justice Perspectives on Fairness Judgments: Comparative Equality or Proportional Equity? 63
3.5.1 Introduction 63
3.5.2 Methods 64
3.5.3 Results and Discussion 66
3.6 General Discussion of Surveys 67
3.7 Appendix B: Chapter Three Tables and Figures 69
Chapter 4: Effects of Justice Sensitivity on Neural Processing of Fairness 94
4.1 Introduction 94
4.2 Methods 98
4.2.1 Participants 98
4.2.2 Dispositional Measures 98
4.2.3 Stimuli and Tasks 99
4.2.4 MRI Acquisition and Processing 101
4.2.5 Data Analysis 101
4.3 Results and Discussion 103

4.3.1. Behavioral Responses 103

4.3.2. Hemodynamic Responses 105

4.4 Appendix C: Chapter Four Tables 114

Chapter 5: Effects of Justice Sensitivity on Spatiotemporal Dynamics of Fairness Processing 139

5.1 Introduction 139

5.2 Methods 144

5.2.1 Participants 144

5.2.2 Dispositional Measures 145

5.2.3 Stimuli and Tasks 145

5.2.4 EEG Acquisition and Processing 146

5.2.5 Data Analysis 147

5.3 Results and Discussion 149

5.3.1 Behavioral Responses 149

5.3.2 Electrophysiological Responses 150

5.4 Appendix D: Chapter Five Tables 152

Chapter 6: General Discussion 162

6.1 Situational Cues of Justice 162

6.2 Neural Systems and Justice Perspectives 164

6.3 Future Directions 166

6.4 Conclusions 168

References 170

Appendix E: Justice Sensitivity Inventory – Short Form 206

Appendix F: Sensitivity to Punishment and Sensitivity to Reward Questionnaire 207

Appendix G: Justice Sensitivity Inventory 209

Appendix H: Interpersonal Reactivity Index – Modified 213

Appendix I: Instructions and questions used in the Knowledge Misrepresentation Task 215

Appendix J: Example stimuli from Three-party Distribution Task used in Chapters 4 and 5.... 220

List of Tables

Table 1. Study 1 Demographics by Group.......... 34

Table 2. Study 1 Dispositional Measures by Group 35

Table 3. Moral Scenarios 36

Table 4. Study 1 Model Parameters for Scenario Responses 41

Table 5. Study 1 Model Parameters for Responses in Random Distribution Task.......... 42

Table 6. Study 1 Model Parameters for Responses in Dots Task.......... 44

Table 7. Study 1 Model Parameters for Freerice Responses 45

Table 8. Study 2 Demographics by Group.......... 69

Table 9. Study 2 Dispositional Measures by Group 70

Table 10. Study 2 Model Parameters for Scenarios.......... 71

Table 11. Study 2 Model Parameters for Distribution Task 73

Table 12. Study 2 Model Parameters for Freerice Responses 75

Table 13. Study 3 Demographic Information 76

Table 14. Study 3 Dispositional Measures 77

Table 15. Study 3 Model Parameters for Scenarios.......... 78

Table 16. Study 3 Model Parameters for TDT 80

Table 17. Study 3 Model Parameters for Freerice Responses 82

Table 18. Study 4 Demographic Information 83

Table 19. Study 4 Dispositional Measures 84

Table 20. Study 4 Model Parameters for Scenario Ratings.......... 85

Table 21. Study 4 Model Parameters for Equity Classification of Distribution Task 87

Table 22. Study 4 Model Parameters for Freerice Responses 89

Table 23. Study 5 Demographic Information 114

Table 24. Study 5 Dispositional Measures 115

Table 25. Study 5 Model Parameters for Scenario Ratings.. 116

Table 26. Study 5 Model Parameters for Responses in Distribution Task 117

Table 27. Study 5 Model Parameters for Responses in Dots Task.. 119

Table 28. Study 5 Model Parameters for Freerice Responses 121

Table 29. Study 5 Whole-Brain Results for OtherUnfair > OtherFair Contrast......................... 122

Table 30. Study 5 Whole-Brain Results for SelfFair > SelfUnfair Contrast 123

Table 31. Study 5 Whole-Brain Results for OtherFairness x SelfFairness Interaction 124

Table 32. Study 5 Whole-Brain Results for OtherFairness x Perspective Interaction............... 125

Table 33. Study 5 Whole-Brain Results for HighReward > LowReward Contrast................... 126

Table 34. Study 5 Whole-Brain Results for RewardType x Perspective Interaction 128

Table 35. Study 6 Demographic Information 152

Table 36. Study 6 Dispositional Measures 153

Table 37. Study 6 Parameters for Linear Model of Scenario Responses.................................. 154

Table 38. Study 6 Model Parameters of Responses in Distribution Task 155

Table 39. Study 6 Model Parameters for Responses during Dots Task..................................... 157

Table 40. Study 6 Model Parameters for Freerice Response.. 159

List of Figures

Figure 1. Stimuli Used in the Random Distribution Task 46

Figure 2. Example Stimuli from the Dots Task 47

Figure 3. SelfFairness by OtherFairness Interaction in the Distribution Task in Study 1 48

Figure 4. Schematic of Task Order in Study 3.. 60

Figure 5. Sample Stimulus for Three-Party Distribution Task 90

Figure 6. SelfFairness by OtherFairness Interaction from Distribution Task in Study 2 91

Figure 7. SelfFairness by OtherFairness Interaction from Distribution Task in Study 3 92

Figure 8. SelfFairness by OtherFairness Interaction from Distribution Task in Study 4 93

Figure 9. Study 5 SelfFairness by OtherFairness Interaction in the Distribution Task 129

Figure 10. Opposing Influences of Justice Sensitivity for Self and Other in Left Striatum 130

Figure 11. Regions Showing SelfFairness x OtherFairness Interaction 131

Figure 12. Self-Oriented Justice Sensitivity and Reduced Interaction Effect in Bilateral Striatum and Central Autonomic Network. 132

Figure 13. Regions showing positive effect of other-oriented justice sensitivity in Dots task... 133

Figure 14. Other-oriented justice sensitivity and increased striatum response for HighReward > LowReward.. 134

Figure 15. Self-Oriented Justice Sensitivity and reduced signal in ACC and left TPJ.............. 135

Figure 16. Self-Oriented Justice Sensitivity and Reduced HighReward > LowReward Response in Bilateral Insula and Hippocampi. 136

Figure 17. Regions with Perspective x Stimulus Type 137

Figure 18. Significant Perspective x Stimulus Type in Anterior Cingulate Cortex................... 138

Figure 19. Study 6 SelfFairness x OtherFairness Interaction 160

Figure 20. ERP traces and scalp plots for SelfFair and SelfUnfair distributions 161

Acknowledgements

This dissertation would not have been possible without the support and guidance of many individuals. First, I want to thank Jean Decety, my mentor and adviser, for giving me the space to make mistakes and teaching me the importance of being good in theory and methods. I also must thank my undergraduate supervisor, Matthew Belmonte, for instilling in me a passion (obsession?) for methodological rigor. My development as a scientist and my approach to research have benefitted immensely from interactions both inside and outside the classroom with many of the faculty in the Integrative Neuroscience program, and I want to express my gratitude to Marc Berman, Jasmin Cloutier, John Cacioppo, Dave Gallo, Dan Margoliash, Greg Norman, Howard Nusbaum, and Steve Shevell. I want to especially thank Marc Berman, Greg Norman, and Steve Shevell for agreeing to serve on my dissertation committee.

I also want to thank the current and former lab mates with whom I have served in these data-drenched trenches. The welcoming sarcasm provided by Eric Porges, Laurie Skelly, and Inbal Bartal played a huge part in helping me find a home in the Social Cognitive Neuroscience Laboratory. My daily experiences preparing and executing this dissertation were deeply enriched by the postdocs with whom I've shared an office: Jason "logic check" Cowell and Cliff "swimmingly" Workman. Finally, to Karen Smith, who was as willing to argue about Doctor Who at Chinatown as to walk to get yummy coffee while discussing the intricacies of multilevel modelling, my sincerest heartfelt thanks.

Abstract

From an early age, people care deeply about justice. However, reactions to injustice are altered by being involved in the situation, and people differ in their sensitivity to justice. Research suggests that dispositional self-focused and other-focused justice sensitivity reflect antisocial and prosocial tendencies. However, it is unclear (1) whether similar behavioral effects can be elicited by brief perspective-taking manipulations and (2) whether these justice orientations are supported by distinct neural systems. Across six studies, participants rated moral scenarios in either a second-person or third-person frame, then played a three-party ultimatum game and completed tasks probing antisocial and prosocial behavior. In Study 1, self-oriented justice sensitivity predicted greater acceptance of distributions after the second-person than third-person frame and, surprisingly, reduced antisocial behavior. Study 2 found no effects of perspective framing, but a negative relation between other-oriented justice sensitivity and acceptance of distributions that were unfair to a third-party. In Study 3, within-participant manipulation of perspective changed rejection rates for distributions which were fair to the self but unfair to another. Study 4 found that proportional equity was a better predictor of behavior than comparative equality. Using functional magnetic resonance imaging, Study 5 found that justice dispositions predicted activity within regions important for saliency processing and executive control. Moreover, these orientations were associated with activity in different networks, suggesting that they arise from distinct processes. Finally, Study 6 failed to find effects of justice sensitivity or perspective framing on electrophysiological measures of neural activity. Overall, these results suggest that perspective framing can influence behavior, though often interacts with justice dispositions. Moreover, justice orientations relate to largely separable neural systems and may prove important for future investigations of social functioning.

Chapter 1: Introduction

Justice is a core component of morality, and involves issues of fairness, equity, and equality (Decety & Yoder, 2017). Justice is also a central concern of human societies and appears to be a fundamental aspect of human nature (Decety & Wheatley, 2015; Vermunt, 2014). Importantly, being personally involved in a situation fundamentally alters one's perception of the event (Baumert, Rothmund, Thomas, Gollwitzer, & Schmitt, 2013). Across similar contexts, people are largely consistent in their attitudes towards justice, are motivated to see others treated fairly, and strive to adhere to justice principles themselves (Baumert & Schmitt, 2016). However, individuals often disagree about the extent to which a given situation deviates from principles of justice or which principles of justice are most applicable. Thus, research indicates that both situational factors and individual sensitivities contribute to judgments about fairness and morality, as well as ethical and unethical behavior (Gino, 2015).

Over the past two decades, social justice researchers have documented stable individual differences in sensitivity to issues of justice in adults (Baumert, Beierlein, et al., 2014; Schmitt, Baumert, Gollwitzer, & Maes, 2010; Schmitt, Neumann, & Montada, 1995). A major finding from this literature is that the link between dispositional justice sensitivity and behavior depends, at least in part, on perspective framing – i.e. whether injustice is directed at the self or another person. Thus, while justice sensitivity captures an individual's general concern for justice, this trait can be divided into two kinds of dimensions. Self-focused processes, such as victim sensitivity, relate specifically to becoming the victim of injustice. Conversely, other-focused processes relate to issues of justice when the potential or actual victim of injustice is another person. Interestingly, these justice "perspectives" appear to represent somewhat competing processes, and actually manifest opposing influences on a wide range of behaviors, including altruistic sharing, moral judgment, and socially transgressive behavior (Decety & Yoder, 2015; Fetchenhauer & Huang, 2004; Gollwitzer, Schmitt, Schalke, Maes, & Baer, 2005). Yet, despite a host of psychological studies on the effects of justice sensitivity, there has been virtually no work attempting to characterize the neural mechanisms by which justice perspectives might contribute to justice-related judgments and behaviors.

This is unfortunate, because multiple lines of research have separately investigated specific neural networks which support social decision-making in general, and these networks likely play a critical role in supporting perceptions of and reactions to injustice. First, moral psychology has established the importance of both outcome and intention when determining someone's moral culpability (Buckholtz et al., 2008; Malle & Guglielmo, 2012; R. M. Miller, Hannikainen, & Cushman, 2014; Young, Cushman, Hauser, & Saxe, 2007). The second line of research comes out of the tradition of neuroeconomics, where researchers have attempted to identify how socioemotional processes contribute to decisions in economic contexts by appealing the domain-general reward and valuation systems (Sanfey, 2007). This endeavor has been largely successful at integrating "social specific" neurocognitive systems such as theory of mind with neural mechanisms for representing value (Ruff & Fehr, 2014). Together, these approaches contribute to a fuller understanding of justice, where evolutionarily deep neurocognitive mechanisms for value learning, reward processing, threat detection, and response selection are coopted to support justice-related behaviors and reasoning which can be applied to dyadic interactions or generalized to three-party contexts (Buckholtz & Marois, 2012; Decety & Yoder, 2017; Shenhav & Greene, 2010)

This dissertation seeks to clarify how justice perspectives (both in terms of situational context and dispositional sensitivity) contribute to justice-related judgments and behaviors by examining spatiotemporal and hemodynamic measures of neural function. It also examines whether previous associations between behaviors and dispositional self-oriented and other-oriented justice sensitivity can also obtain from situational cues towards self-focused and other-focused processing. Previous investigations into justice processing have focused on motivation for justice as a personality trait and the specific links between justice perspectives and attention allocation, ambiguous scene interpretation, prosocial behavior, antisocial behavior, and moral judgment (Baumert, Rothmund, et al., 2013; Decety & Yoder, 2015; Hafer & Begue, 2005; Tyler, 1994). However, these investigations have been almost exclusively limited to behavioral studies. Thus, very little is known about how individuals differences in justice motivation relate to neural activity (Alexopoulos, Pfabigan, Göschl, Bauer, & Fischmeister, 2013; Boksem & De Cremer, 2010; Yoder & Decety, 2014a, 2014b). Establishing these links is crucial, because many of the putative components of self-oriented and other-oriented justice sensitivity, such as enhanced information processing, lower activation thresholds, and motivation to re-establish justice, have clear predictions for their underlying neural circuitry. For instance, enhanced attention towards others' potentially exploitative intentions, a core facet of self-focused justice sensitivity (Gollwitzer, Rothmund, & Süssenbach, 2013), should involve attention allocation and theory of mind, and thus manifest as increased hemodynamic activity within the dorsal attention network (Corbetta & Shulman, 2002; Dosenbach et al., 2007) such as inferior parietal lobe, medial prefrontal cortex (mPFC), and dorsolateral prefrontal cortex (dlPFC). Similarly, since dispositional other-oriented justice motivation involves increased automatic attention capture by injustice-related information (Baumert, Gollwitzer, Staubach, & Schmitt, 2011), these dispositions should also influence regions implicated in saliency processing such as anterior insula (aINS) and dorsal anterior cingulate cortex (dACC), as well as specific event-related potentials which index attention capture, such as N200 and the related medial frontal negativity (Hajcak, Holroyd, Moser, & Simons, 2005; Hajcak, MacNamara, & Olvet, 2010).

1.1 Justice as a Universal Motive

There have been several attempts to unify the immense diversity in moral and ethical codes seen across cultures. Some approaches argue for a small number of distinct systems which exist in all individuals and societies but are expressed to varying degrees (Gilligan, 1982; Graham et al., 2011; Shweder, Much, Mahapatra, & Park, 1997). Others suggest that moral issues belong to a unitary domain (Kohlberg, 1968; Nucci & Turiel, 1978; Smetana, 2013). Despite their disagreements, justice principles always play a central role in structuring at least some aspects of the moral domain. For instance, justice is an essential component of the ethic of autonomy in the "Big Three" (Shweder & Haidt, 1993) and is the driving force in the Fairness/Reciprocity moral foundation (Graham et al., 2011).

Moreover, several investigations into non-human animals indicate that precursors to the justice motive may have evolved because of ecological pressures to stabilize cooperation within groups. For instance, several primate species demonstrate inequity aversion and respond negatively when they receive less reward than a social partner (Brosnan, 2013). Studies in macaques have been crucial for mapping out neural populations which encode differences in outcomes both with respect to previous outcomes experienced by the individual and concurrent rewards obtained by the individual and a social partners. Such regions include the lateral prefrontal cortex (Hosokawa & Watanabe, 2015), dACC (Haroush & Williams, 2015), and striatum (Báez-Mendoza, Harris, & Schultz, 2013; Báez-Mendoza, van Coeverden, & Schultz, 2016). Interestingly, among the primates studied to date, inequity aversion has only been observed in those species whose social structure requires them to regularly cooperate with non-kin (Brosnan, 2013; Brosnan & Bshary, 2016), suggesting that negative responses to inequitable distributions might be fundamental for motivating individuals to adhere to mutually beneficial outcomes in groups of genetically unrelated individuals.

While such comparative approaches can provide important unique insights into the ultimate causes of justice motivation, including guideposts towards identifying essential neural systems, it is important to note that there is little evidence that any non-human species demonstrate more than basic elements of justice motivation (Thierry, 2000). Indeed, while justice appears to be a basic human impulse (Tyler, 1994), mature justice motivation involves appeals to abstract principles which are maintained by social and cultural institutions (Brosnan & Bshary, 2016; Wilson, 2012). Thus, the proximate mechanisms underlying the justice motive, including affective and cognitive processes, must also be investigated. Research indicates that people generally condemn deviations from principles of justice as morally wrong, recommend punishment for others when they violate such principles, and strive themselves to behave in accordance with justice (Decety & Wheatley, 2015; Vermunt, 2014). A growing number of studies find that these capacities emerge early in development, and that even in infancy there is a general tendency to identify variations in fairness and prefer or approach justice and dislike or avoid injustice (Cowell & Decety, 2015b; Hamlin, 2014; Jensen, Vaish, & Schmidt, 2014; Schmidt & Sommerville, 2011). There is also some tentative evidence that individuals find some aspects of justice, such as fairness and cooperation, inherently rewarding (Fehr & Camerer, 2007; Tabibnia, Satpute, & Lieberman, 2008).

At the same time, individuals sometimes do engage in behaviors that violate principles of justice (Gino, 2015; Moore & Gino, 2013), and a growing number of studies highlight the importance of prefrontal-amygdalar-striatal interactions when engaging in unethical behaviors such as dishonesty and deception (Abe, 2011; Garrett, Lazzaro, Ariely, & Sharot, 2016; Sun, Chan, Hu, Wang, & Lee, 2015). Thus, while a neuroscience perspective of justice sensitivity seeks to identify the neural mechanisms underlying perceptions, interpretations, and judgments about issues of fairness, equity, and rights, this dissertation also seeks to characterize how these factors are translated into behaviors that may or may not accord with principles of justice. Some theories have begun to address this, such as work on the Third Party Punishment network, which has been proposed as the neurobiological underpinnings of the remarkably human capacity for creating cultural institutions dedicated to enforcing social norms (Buckholtz & Marois, 2012). According to this model, cultural institutions of justice arise from strong evolutionary and social pressures to enforce cooperation, which in turn rely on more basic neurocognitive systems for value learning, decision-making, and perspective-taking. For example, in one study participants were asked to imagine themselves as members of jury, and then recommend sentences (i.e. punishments) for crimes which varied in outcome severity and responsibility of the perpetrator (Buckholtz et al., 2008). According to this model, the amygdala encodes affective arousal related to the harm of others while the temporoparietal junction (TPJ) or posterior superior temporal sulcus (pSTS) encodes mental state information. This information is then integrated in mPFC/dACC and decisions about culpability and sanction recommendations are driven by dlPFC (Treadway et al., 2014). If individuals evaluate their own behavior in a similar fashion, then variation in such a network may play a critical role in giving rise to individual differences in justice sensitivity, but this remains a largely unanswered question (Yoder & Decety, 2014b).

1.2 The Economics of Justice

One fruitful line of research into justice motivation has utilized economic games to examine how individuals make decisions about and react to equality and inequality. These approaches have been successful in large part because they allow a great deal of experimental control over objective measures of outcome equality. For example, the Ultimatum Game (Güth, Schmittberger, & Schwarze, 1982) has been widely used to assess perceptions of and responses to unfairness (Sanfey, 2007). In the Ultimatum Game (UG), one person plays the role of the Proposer, and decides how to divide some resource (e.g. typically money, though stickers or candies are often used with children) between themselves and a second player, the Responder. The Responder then decides to accept or reject the offer. If the offer is rejected, both players get nothing. If people were motivated solely by self-interest, then Proposers should offer very little money and Responders should accept any nonzero offer. However, this is rarely the case, with many Proposers choosing to offer close to 50%, and most Responders rejecting offers of less than one third of the initial amount (Gabay, Radua, Kempton, & Mehta, 2014; Güth et al., 1982; Sanfey, 2007). Thus, it appears that individuals in both roles are aware of some sort of "equal division" expectation, and are motivated to pursue this goal. Importantly, this appears to be a *social* justice rule, because unfair offers are accepted at a higher rate when made by a computer than by another person (e.g. Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003).

Importantly, UG rejections are also sometimes conceptualized as altruistic punishment (Fehr & Fischbacher, 2004), because individuals are willing give up their potential small reward to prevent the Proposer for violating the fairness norm. Such costly punishment exists across societies, though there are some variations between cultures (Cowell et al., 2016; Henrich et al., 2006, 2010). While, individuals vary in the amount they are willing to accept (Gabay et al., 2014), ultimatum-style games have proven to be an effective tool for investigating the effects of perspective-taking because rejection decisions are reliably altered by changing a participant's role. Such investigations have primarily adopted one of two approaches. In the first, participants play the role of Responder in two successive games, once for themselves, and once for another person (e.g. Civai, Crescentini, Rustichini, & Rumiati, 2012; Corradi-Dell'Acqua, Civai, Rumiati, & Fink, 2013). This manipulation allows researchers to examine unique contributions of self-interest above and beyond general concern with justice. The second, though slightly more complicated approach, introduces a third player (e.g., Alexopoulos, Pfabigan, Lamm, Bauer, & Fischmeister, 2012). In such three-party games, the Proposer makes a single offer to split their initial endowment between two other people, one Responder and one Observer (Dawes et al., 2012). While the Responder plays the same role as in the classical UG, the Observer is not given an opportunity to influence the outcome and must passively observe the offer and decision. This structure allows for distributions which pit justice motivations (e.g. distributional equality, avoiding selfish actions) against self-interested motivations (e.g. maximizing personal payoff).

Since ultimatum-style games provide good experimental control over objective equality, they have been repeatedly used to investigate the neural systems which underlie perceptions and reactions to unfairness (Ruff & Fehr, 2014; Sanfey, 2007). For instance, a recent meta-analysis of 20 functional magnetic resonance imaging (fMRI) studies indicated that when Responders are presented with unfair compared to fair offers, they reliably show greater hemodynamic responses in a distributed set of regions, including bilateral aINS and ventrolateral prefrontal cortex (vlPFC), dACC, dorsal mPFC, dlPFC, inferior parietal lobule (IPL), and precuneus. Conversely, comparing fair to unfair offers consistently elicits responses in posterior insula, ventromedial prefrontal cortex (vmPFC), inferior temporal gyrus, and posterior cingulate cortex (PCC), as well as a slightly more dorsal region of the precuneus (Feng, Luo, & Krueger, 2015). There is also strong evidence that these activation patterns reflect the recruitment of domain-general value representation systems that additionally include the amygdala and striatum (Gospic et al., 2011;

Tabibnia et al., 2008). Many of these same regions were implicated in the Third Person Punishment network discussed above, providing converging evidence that the neural underpinnings of justice build on a core valuation circuitry which tracks the reward expectancies of both monetary and social issues (Buckholtz & Marois, 2012; Ruff & Fehr, 2014).

Importantly, activity within these networks is altered by one's perspective. For instance, dACC and vmPFC show greater response to unfair offers when participants are deciding for themselves, but not when they make decisions on behalf of a stranger (Civai et al., 2012; Corradi-Dell'Acqua et al., 2013). Conversely, aINS response appears to respond to unfairness (deviation from even split) in general, and shows a similar activity pattern for unfair > fair regardless of the participant's perspective. This is also true during three-party random distribution games, where allocations arise from a computer (Dawes et al., 2012), suggesting that economic fairness processing in aINS does not require that distributions arise from a social agent. Moreover, aINS response uniquely correlates with rejection rates for unfair offers, as well as generosity outside the scanner, suggesting that aINS response may index an individual's general concern with social norms, including fairness and justice (Civai et al., 2012). Some have also argued that aINS activity may index the interoceptive experience of negative emotions associated with unfairness (Sanfey et al., 2003)

However, aINS plays an important role in a variety of cognitive and emotional tasks (Craig, 2009; Menon & Uddin, 2010) that are not necessarily directly linked to justice principles. In this way, aINS response to unfair offers might also reflect a deviation from expected outcomes, such as equal distributions (Civai, 2013). This fits with the more general function played by aINS in the salience network, where aINS and dACC serve as core input and output nodes in a distributed neural circuit which functions to coordinate behavioral responses to motivationally relevant external and internal stimuli (Harsay, Spaan, Wijnen, & Ridderinkhof, 2012; Seeley et al., 2007; Shackman et al., 2011).

Unfair offers in the ultimatum game have also been linked to electrophysiological correlates of ACC activity. Several different tasks reliably elicit a negatively going scalp potential over frontal midline sites around 200-350 ms (often called the medial frontal negativity; MFN) after negative outcomes, especially unexpected negative outcomes, including financial losses and committing an error (Gehring & Willoughby, 2002; Hajcak, Moser, Yeung, & Simons, 2005; Nieuwenhuys, 2012). MFN amplitude is enhanced for the presentation of unfair compared to fair offers in ultimatum games, and this effect is increased by dispositional concern for fairness (Boksem & De Cremer, 2010) or experiencing social victimization (Qu, Wang, & Huang, 2013). Importantly, MFN amplitudes are enhanced by negative feedback not only to one's own actions, but also when observing others (Koban, Pourtois, Bediou, & Vuilleumier, 2012). Moreover, this effect is enhanced in cooperative contexts and reduced in competitive contexts, suggesting that social relationships may influence early stages of justice-related information processing.

While economic games do provide good experimental control, the simplicity of economic games is a double-edged sword. Insights neuroeconomics are crucial to understanding the neural underpinnings of fairness judgments, yet many everyday justice issues are not as readily apparent. Instead, social interactions are complex dynamic processes, and the relevance of specific justice principles often depends on how an individual construes a situation (Baumert & Schmitt, 2009). Moreover, not all justice principles apply to all situations. For instance, while justice would argue that medicine should be distributed first to those that are sick rather than those that are healthy, clean water should be available to everyone. Thus, while context is a crucial factor at determining which principles of justice are relevant to a particular situation (Carnes, Lickel, & Janoff-Bulman, 2015), differences in perspective (i.e. how one frames a situation), are an equally crucial factor in how a situations is perceived.

1.3 Justice as a Personality Trait

People differ in the frequency with which they perceive injustice as well as the extent to which they are motivated to respond to unjust acts (Schmitt et al., 2010; Vermunt, 2014). Over the past two decades, social justice researchers have documented stable individual differences in sensitivity to issues of justice in adults (Baumert, Beierlein, et al., 2014; Schmitt et al., 2010, 1995), and more recently adolescents and children (Bondü & Elsner, 2015). Though these dispositions are quite stable in adulthood (Dar & Resh, 2001; Lovas & Wolt, 2002; Schmitt, Gollwitzer, Maes, & Arbach, 2005), they may be experience-expectant, and research suggests that they are influenced by previous experiences with justice and injustice (Gollwitzer, Süssenbach, & Hannuschke, 2015; Wijn & Bos, 2010) and are likely responsive to parental transmission of values (Cowell & Decety, 2015a).

The earliest concepts of justice motivation built on social resource theory (Foa & Foa, 1974), which posits that social interactions are best understood as "allocation events," in which one individual allocates one of six types of resources (e.g. money, information) to another. In this conception, the justice motive was a motivation to perform allocations in a way that balanced self-interest and other-interest (Vermunt, 2014). This contrasts with just world theories which argue that concern for others depends on their deservingness (Lerner, 1980), as well as deonance theories which argue that individuals are morally obligated to help others, regardless of the cost to self (Folger, 2011).

One of the most widely used measures of dispositions regarding justice is the Justice Sensitivity Inventory (JSI), a self-report scale ranging from 8 to 40 items depending on the version used (Baumert, Beierlein, et al., 2014; Schmitt et al., 2010, 2005). In contrast to the Belief in a Just World scale (Hafer & Begue, 2005), which assesses an individual's need for justice to be an organizing principle of the world, justice sensitivity captures an individual's personal commitment to justice, and measures someone's propensity to detect and respond to injustice from four perspectives (Baumert, Halmburger, & Schmitt, 2013a). The Victim perspective measures how sensitive an individual is to being exploited by others. Observer, Beneficiary, and Perpetrator perspectives, as their names suggest, assess one's sensitivity to injustice when one is a neutral observer, when one benefits from injustice directed at others, or when one is the cause of an injustice.

A common use of the JSI is to collapse JS-Observer, JS-Beneficiary, and JS-Perpetrator into a single factor representing dispositional other-oriented sensitivity and contrast this against JS-Victim as a self-oriented disposition (Fetchenhauer & Huang, 2004; Lotz, Baumert, Schlösser, Gresser, & Fetchenhauer, 2011; Lotz, Schlösser, Cain, & Fetchenhauer, 2013). In fact, these other-oriented justice dispositions often load together and are very highly inter-correlated (Baumert, Rothmund, et al., 2013; Fetchenhauer & Huang, 2004; Lotz et al., 2011). While self-oriented and other-oriented justice sensitivity perspectives confer certain information processing benefits for concepts related to justice, they often represent competing prosocial and antisocial motivations (Decety & Yoder, 2015; Gollwitzer et al., 2013).

In terms of information processing, both perspectives have been linked to lower activation potential and increased availability of concepts related to injustice, as compared to neutral or negative controls. For instance, several studies have demonstrated that when individuals with high other-oriented sensitivity watch a video depicting an unjust situation (e.g. mocking someone because of their religion), they interpret a subsequent ambiguous interaction as more unjust (Baumert et al., 2011; Baumert & Schmitt, 2009). Similarly, individuals with higher self-oriented sensitivity are faster to complete sentence fragments in first-person narratives when they are semantically related to justice or injustice, as compared to neutral words (Baumert, Otto, Thomas, Bobocel, & Schmitt, 2012). Finally, both perspectives are positively correlated with d' statistics for whether or not a word was novel, specifically for injustice-related words, suggesting that self and other orientations may share enhanced cognitive elaboration of justice and injustice concepts (Baumert et al., 2011, 2012).

However, these orientations produce vastly differ consequences on behavior and judgment. In the context of economic games, individuals with high other-oriented sensitivity show increased generosity and cooperation (Baumert, Schlösser, & Schmitt, 2014; Edele, Dziobek, & Keller, 2013). Moreover, these individuals do not significantly alter their behavior and continue to behave prosocially in the face of strong evidence that others players will behave selfishly (Lotz et al., 2011) or when situational factors provide an excuse to behave less altruistically (Lotz et al., 2013). In contrast, individuals with high self-oriented sensitivity reject more offers in the ultimatum game (Baumert, Schlösser, et al., 2014), are less generous (Fetchenhauer & Huang, 2004), and disengage from cooperative trust games in response to even slight cues that others might also defect (Gollwitzer, Rothmund, Pfeiffer, & Ensenbach, 2009).

Fewer studies have examined dispositional sensitivity in non-economic settings. One study found that higher other-oriented sensitivity increases the likelihood that someone will intervene to stop a theft (Baumert, Halmburger, et al., 2013a). Conversely, higher self-oriented sensitivity reduces the odds that a person will engage in morally courageous behavior, such as providing a name and home address to sign a petition against a right-wing extremist political party (Kayser, Greitemeyer, Fischer, & Frey, 2010). Individuals with higher self-oriented justice sensitivity also report a greater number of previous transgressions (Gollwitzer et al., 2005). Moreover, self and other orientations exert opposing influences on judgments about others' hypothetical social and moral transgressions, with other-orientated scores predicting greater condemnation, but self-oriented scores predicting higher permissibility ratings (Decety & Yoder, 2015; Gollwitzer et al., 2005).

Given these diverse findings, it is not surprising that recent conceptions of justice sensitivity as a personality trait have argued that self and other sensitivities reflect antisocial and prosocial motivations, respectively. In other words, while other-oriented sensitivity represents a genuine prosocial motivation and unconditional concern for justice as a moral principle, self-oriented sensitivity reflects self-interested concern for oneself, potentially motivated by a fear of being victimized by others (Baumert, Rothmund, et al., 2013; Gollwitzer et al., 2013, 2015; Süssenbach & Gollwitzer, 2014). Thus, the reason individuals with higher other-oriented sensitivity behave prosocially, such as when they persist in cooperative behaviors in situations where they know they may be taken advantage of, is because they have a stronger sense of moral identity and desire to behave in accordance with justice principles (Gollwitzer et al., 2009). At the same time, highly victim sensitive individuals behave in ways that minimize the chances of being exploited, because they systematically underestimate the trustworthiness of others (Gollwitzer, Rothmund, Alt, & Jekel, 2012; Gollwitzer et al., 2013).

Overall, given their relations with moral decision-making and moral behavior, both in the context of economic games and broader everyday behaviors, JSI perspectives provide a good index of individual differences in justice perspectives. Thus, they should also provide important insights into how dispositional differences in justice orientations are related to variations in neural systems across a wide range of tasks. However, few studies to date have attempted this. One fMRI study found that other-oriented sensitivity was positively correlated with hemodynamic differentiation between morally good and bad actions in right dlPFC and dmPFC, providing some preliminary support for the idea that individuals that are more sensitive to justice as a moral principle may rely more on prefrontal circuits to evaluate interpersonal interactions (Yoder & Decety, 2014b).

Based on the literature reviewed above, this dissertation sought to identify the influence of justice perspectives on moral judgments and behaviors, both at the behavioral and neural level. People do value justice and are motivated to behave in accordance with justice principles, but when one is personally involved in a situation, the motivation to see justice enacted in the world may be more heavily influences by the motivations to benefit the self and avoid self-negative outcomes (Decety & Yoder, 2017). While both theory and research argue that self-oriented and other-oriented justice sensitivity are separable (Baumert, Halmburger, & Schmitt, 2013b; Baumert & Schmitt, 2016; Decety & Yoder, 2015, 2017, Gollwitzer et al., 2009, 2005), there is a paucity of research examining what determines when self-oriented or other-oriented sensitivity dominates information processing at any given time. There has been a suggestion that justice sensitivity dispositions "become activated by relevant situational cues" (Baumert, Rothmund, et al., 2013, p. 169); however, it is unclear whether this means that stable inter-individual differences in justice sensitivity render some individuals more or less likely to respond to these cues, or rather that all individuals can be moved towards other-oriented or self-oriented processing by situational factors. Moreover, justice dispositions do appear to be relatively stable across contexts and produce downstream consequences on justice-related information processing and behavior. What has not been well-studied is the extent to which self-focused and other-focused justice processes are supported by the nervous system. In other words, it seems plausible that self-focused and other-focused justice perspectives may rely on distinct neural networks, but this is an open question. This dissertation proposes to begin to address these gaps in the literature by focusing on two specific questions:

Question 1 (*Chapters 2-3*): *Do situational cues for self-focused and other-focused justice influence behavior in a similar fashion to justice dispositions independent of those dispositions?*

Question 2 (*Chapters 4-5*): *Which neural systems are differentially responsive to other-focused and self-focused justice processing?*

Chapter 2: Behavioral Effects of Justice Sensitivity and Perspective

2.1 Introduction

Investigations into individuals' reactions to injustice, whether directed at themselves or third parties, highlights how justice-related concerns (e.g. fairness, equity, equality) influence social decision-making (Baumert & Schmitt, 2016; Decety & Yoder, 2017; Clara Sabbagh & Schmitt, 2016). Moreover, understanding the factors that contribute to decisions to engage in both antisocial and prosocial behaviors have important implications for societal functioning (Gino, 2015; Glenn & Raine, 2014). All of these behaviors show individual differences, and while personality factor are certainly important, social psychology has long highlighted the importance of situational cues (Ross & Nisbett, 1991). The purpose of Study 1 was to explore whether situational cues to experiencing or observing injustice influence justice-related behaviors, and whether shifts in these perspectives impact behavior by interacting with individual dispositions in justice sensitivity. Specifically, Study 1 investigated the effects of perspective framing and justice sensitivity on moral judgment, economic decision-making, dishonesty, and helping.

Much of previous research into moral cognition has been dominated by a focus on sacrificial dilemmas (Kahane, 2015). Perhaps the most famous such dilemma is the Trolley Problem (Foot, 1967; Thomson, 1985), in which participants are asked to decide if it is morally permissible to pull a lever and divert a runaway trolley onto a different track, saving the lives of five strangers, but consequently leading to the death of one stranger. Ostensibly, pulling the lever represents the utilitarian choice, while not pulling the lever is indicative of deontological ethics. Such dilemmas have been used to support dual-process accounts of moral reasoning in which deontological rules depend primarily on emotion, and utilitarian deliberations depend on cognition (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). Sacrificial moral dilemmas have also been used to claim specific kinds of emotional deficits in populations with lesions (Michael Koenigs et al., 2007; Taber-Thomas et al., 2014) or psychopathy (Michael Koenigs, Kruepke, Zeier, & Newman, 2012; Rosas & Koenigs, 2014). However, this approach has been challenged for its weak external validity, with critics claiming that neither choice presents a true utilitarian option, that the dilemmas are not moral dilemmas, and that they are unlikely to reflect how moral decision-making actually occurs for most people during their everyday lives (Kahane, Everett, Earp, Farias, & Savulescu, 2015).

An alternative approach has been to examine how domain-general systems contribute to moral judgments in more commonplace, everyday situations (Decety & Yoder, 2015; Shenhav & Greene, 2010). This approach has successfully been used in cognitive neuroscience to address questions about the various roles of perspective-taking, emotional reactivity, and executive functioning in moral cognition (Avram et al., 2014; Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006; Yoder & Decety, 2014a, 2014b). Similar everyday scenarios have also been used to assess differences between typical and psychiatric populations (Yoder, Harenski, Kiehl, & Decety, 2015; Yoder, Lahey, & Decety, 2016), and in developmental neuroscience to probe the emergence of various aspects of moral reasoning (Cowell & Decety, 2015a, 2015b).

One relatively recent notable theoretical addition to the field of moral psychology is the distinction between proscriptive and prescriptive moral motivations. This work builds on a theoretical model of fundamental systems which support behavioral approach and behavioral avoidance (Gray, 1990). In this model, the Behavioral Activation System (BAS) is sensitive to rewards and facilitates approach towards appetitive stimuli. Conversely, the Behavioral Inhibition System (BIS) is sensitive to punishment and supports the avoidance of aversive stimuli. These distinctions have been further developed in personality psychology, such as research utilizing questionnaires to assess BIS and BAS (Carver & White, 1994; Cooper & Gomez, 2008) and promotion vs. prevention in regulatory focus theory (Higgins, 1997). In the Moral Motives framework (Janoff-Bulman, Sheikh, & Baldacci, 2008), approach motivation and avoidance motivation support distinct types of morality. Approach motivation supports prescriptive morality, which covers morally obligatory actions, while proscriptive morality, which is concerned primarily with forbidden actions, is subserved by avoidance motivation (Janoff-Bulman, Sheikh, & Hepp, 2009; Sheikh & Janoff-Bulman, 2010).

Importantly, moral regulation theories also interface with a self-focused and other-focused distinction. Thus, there exist self-focused approach and avoidance motivations, as well as other-focused approach and avoidance mechanisms. More recent theoretical formulations also argue for a group-level collective focus (Janoff-Bulman & Carnes, 2013). However, the specific relations between self-focused and other-focused dispositional justice sensitivity and approach-based or avoidance-based motivational systems are currently unknown. Moreover, a better understanding of such associations can also shed light on important philosophical questions. For instance, one unresolved debate is how best to explain supererogatory actions, i.e. actions which are morally praiseworthy, but not morally required (Dorsey, 2013). Neither purely utilitarian nor deontological formulations can account for actions of this sort (Ferry, 2013), but examining asymmetries in self-focused and other-focused justice sensitivity might offer a resolution. Dispositional justice sensitivity has been linked to interpretation biases (Baumert & Schmitt, 2009), suggesting that some individuals may perceive prescriptive norms as applicable in more situations or as carrying more moral force. Thus, the scenarios employed here include both proscriptive and prescriptive violations.

It was hypothesized that scenarios negatively affecting the self would be judged to be less permissible, and that this would be enhanced for proscriptive compared to prescriptive outcomes. Again, other-focused justice sensitivity has been argued to reflect prosocial concern for the welfare of others, while self-focused sensitivity appears to capture, at least in part, motivation to avoid exploitation (Süssenbach & Gollwitzer, 2014). Thus, other-focused and self-focused justice dispositions were hypothesized to predict specific condemnation of prescriptive and proscriptive transgressions, respectively.

The scenarios used in Study 1 also served to manipulate participant's perspectives. By manipulating whether the potential victim of an event was the participant or a third-party, the moral judgments themselves were expected to induce a self-focused or other-focused mindset. Early investigations into victim sensitivity proposed a cluster of attributes which define individuals with victim sensitivity: greater frequency of experienced injustice, more intense emotional (e.g. anger, moral outrage, guilt) reactions to injustice, rumination on unjust events, and a motivation to restore justice, potentially by punishing their victimizer (Schmitt et al., 1995). Later studies argued that these individuals also possess a lower activation threshold for justice- and injustice-related conceptions, and so have a "chronic accessibility" of such concepts (Baumert & Schmitt, 2009). Thus, it was expected that self-oriented and other-oriented justice sensitivity dispositions would interact with this perspective framing, with justice sensitivity predicting which individuals would show a greater impact on behavior.

As discussed in *Chapter 1*, decisions to reject unfair offers during ultimatum-style games can serve as a measure of an individual's commitment to fairness norms (Henrich et al., 2006). Moreover, by including a third party, it is possible to independently manipulate fairness for oneself and others. Here, I predicted that individuals who were self-focused, either as a result of scenario perspective framing or because of high trait self-oriented justice sensitivity, would subsequently focus more on their own payoffs and thus be more likely to accept allocations which were fair to themselves, regardless of whether those offers were fair to the other. Conversely, other-focused participants were expected to take the fairness of others into account, and thus reject more offers which were unfair to the others.

Finally, justice sensitivity dispositions have been argued to reflect prosocial and antisocial motivations (Süssenbach & Gollwitzer, 2014). For instance, self-oriented justice sensitivity has been linked to higher self-reported transgressions (Gollwitzer et al., 2005), while other-oriented justice sensitivity predicts whether individuals will intervene to prevent theft (Baumert, Halmburger, et al., 2013a). However, because previous studies have investigated antisocial or prosocial behavior independently, rather than both together, it is currently unknown whether these asymmetrical effects exist within the same individuals. To address this, Study 1 included tasks which provided the opportunity for participants to behave antisocially (earn more money for dishonest responses) and prosocially (spend time answering questions to donate food). I predicted that self-focused individuals would show greater antisocial behavior, coincident with reduced prosocial behavior. Conversely, other-focused individuals were expected to show the opposite pattern.

2.2 Methods

2.2.1 Participants

Study 1 was conducted at the Chicago Research Lab in downtown Chicago, where 100 adults were recruited from the local participant pool. All participants provided informed written consent and were paid \$5.00 plus whatever they earned during the Dots Task (*M* = \$5.48). Three participants failed to provide complete questionnaire information, leaving a starting sample of 97 adults (see Table 1 for demographic information). All materials and tasks were approved by the Institutional Review Board at the University of Chicago.

2.2.2. Dispositional Measures

After consent, participants completed several questionnaires on a computer and provided demographic information, including age, gender, educational level, and income. The short-form version of the Justice Sensitivity Inventory (Appendix E) was used to obtain a measure of individual's dispositional justice sensitivity (Baumert, Beierlein, et al., 2014). Participants rated how well each of 8 statements described them using a scale from 0 (not at all) to 6 (exactly). Victim sensitivity was taken as an index of self-oriented justice sensitivity (JS-Self), while other-oriented justice perspectives were combined to obtain a single measure of other-oriented justice sensitivity (JS-Other). The 24-item Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ; Appendix F; Cooper & Gomez, 2008) was used to index reward sensitivity and punishment sensitivity as proxies for approach and avoidance motivation. Group averages and reliability for each measure are shown in Table 2.

2.2.3. Stimuli and Tasks

All stimuli (except for the Freerice website [see below]) were presented via E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, USA). Participants were seated in front of computers in a room with four cubicles at the Chicago Research Lab (CRL). As is common in the CRL, multiple participants could begin and finish the study in the same room, so some participants began the study while another participant was in the middle of a task. An experimenter sat at a desk behind the participants and was present throughout the duration of testing.

Moral Framing Scenarios. After completing questions, participants read 10 scenarios adapted from a previous study of justice sensitivity and moral decision-making (Decety & Yoder, 2015). Each scenario consisted of three sentences and described a potential everyday dyadic interaction. As in the prior study, participants were asked how morally permissible it would be for one person to behave in a particular way and indicated their answer using a 7-point scale (1 = Completely, 4 = Somewhat, 7 = Not at all). Half of the scenarios involved proscriptive violations of individuals engaging in a morally questionable action (e.g. cheating on a test, pushing a stranger). The other five scenarios depicted prescriptive violations of individuals failing to engage in a prosocial behavior (e.g. not giving up one's seat for a person on crutches, not helping someone pick up spilled groceries). Participants were randomly assigned to either the Second-Person or Third-Person framing group. Because of technical problems, behavioral responses from seven participants were unavailable. In the second-person group (n = 46), scenarios were worded such that the person affected by the interaction was "you." In the third-person group (n = 44), scenarios were written in the third-person, and the affected party was replaced with "someone." Scenarios were presented in randomized order, and participants completed the task at their own pace (See Table 3 for full text of scenarios). Participants responded by pressing the appropriate number on a computer keyboard.

Distributional Fairness. After rating moral scenarios, participants next completed 40 trials of a Random Distribution Task (RDT). In the RDT, $12 is divided between the participant and two strangers (Figure 1). After the distribution, participants decide whether to accept or reject the distribution. Here, "accept" means that each party keeps the money in front of them. In contrast, "reject" means no one gets any money. The random nature of the distributions builds on previous investigations into fairness judgments (Civai, Corradi-Dell'Acqua, Gamer, & Rumiati, 2010; Sanfey et al., 2003). While these studies show that computer-generated unfair distributions are accepted at lower rates that human-generated offers, the rejection rate is non-zero. Moreover, by removing an agent that is responsible for the distribution, rejection decisions in RDT should be more related to a preference for fairness, rather than reciprocity or punishment motivations (Crockett, Clark, Lieberman, Tabibnia, & Robbins, 2010; Tabibnia et al., 2008)

Distributions were displayed for three seconds before "Accept or Reject?" appeared in the center of the screen to prompt participants to make a response. If participants did not respond within six seconds, a screen appeared warning them to respond more quickly. The monetary distribution remained on screen throughout the response period. Trials were separated by a fixation cross lasting three seconds. Individual allocations were either $2, $4, $5, or $6. This arrangement allowed for an independent manipulation of equality for the self and the equality of the other two players. For instance, if Player 1 is allocated $2, Player 2 is offered $6, and the participant is offered $4, this was coded as SelfUnfair (the participant received less than another player) and OtherUnfair (Player 2 received more than Player 1). In this 2 Other (Unfair|Fair) x 2 Self (Unfair|Fair) design, participants were presented with 10 distributions of each type, for a total of 40 distributions.

Antisocial Behavior. The Dots Task has been repeatedly used to assess unethical behavior and dishonesty (Gino & Ariely, 2012; Mazar & Zhong, 2010). Participants were first informed that they would be performing a perceptual accuracy task. Stimulus displays contained a square that was bisected by a diagonal line. For each trial, 20 red dots appeared within the square and participants had to indicate whether there were more dots on the left or ride side of the diagonal line (see Figure 2 for example stimuli). The dots remained on the screen for one second, and then a response screen appeared with the words "Left" and "Right," prompting participants to indicate their response by pressing a key on the left side ("A") or right side ("L") of the keyboard. When participants pressed one of the response keys, a box appeared around the corresponding word to show that the response had been recorded. The response screen always remained visible for two seconds. If participants did not respond within two seconds, the trial was marked as a "Miss" and excluded from analysis.

After introduction to the task structure, participants completed 50 practice trials. Once the practice trials were completed, the payment rules were revealed. Participants were informed that they would earn money for every response, but that "Right" side responses were rewarded at ten times the rate of "Left" side responses (i.e. 1 cent or 10 cents). The payment instructions also stated that the task itself would consist of 100 trials and reported how much the participant would have earned based on their responses during the 50 practice trials. Importantly, participants were told that their responses would be rewarded regardless of accuracy, but asked to respond as accurately as possible because their "results will be used to design future studies about perceptual accuracy." The task was broken up into four runs. In keeping with previous use of the Dots task (Mazar & Zhong, 2010), each run contained 16 LowReward trials (more dots on left, individual reward $0.01) and 9 HighReward trials (more dots on right, individual reward $0.10). Thus, participants could earn $4.24 by responding with 100% accuracy, or up to $10.00 by selecting the HighReward response on every trial.

Prosocial Behavior. Prosocial behavior was assessed by giving participants the opportunity to spend time playing Freerice (http://freerice.com). The Freerice website asks participants multiple choice vocabulary questions, with difficulty slowly increasing after each correct response. For every correct answer, 10 grains of rice are donated to the World Food Programme (http://wfp.org). Once participants had finished the Dots task, but before learning about the Freerice task, they were paid and signed a payment receipt form. After this payment procedure, participants were then informed about the website as I or a research assistant loaded the Freerice website (e.g. "Our lab has found this cool website where you can answer vocabulary questions and they donate rice to developing countries"). It was stressed that answering questions was entirely voluntary, that participants could not earn more money, and that participants were free to leave immediately if they so desired. Thus, spending time playing Freerice was taken to reflect prosocial behavior, because participants were explicitly told that they could not earn more money, yet their behavior produced a real-world benefit to strangers.

2.2.4. Data Analysis

All analyses were conducted in R version 3.2.3 (R Core Team, 2015). A composite socioeconomic status (SES) score was created as the average of normalized education and income. Participants' responses within each task were analyzed using hierarchical generalized linear multilievel models (HGLMMs). Ordinal multilevel regression was employed to examine individual scenario ratings by fitting cumulative link mixed models (CLMMs) using the 'ordinal' package (Christensen, 2015). Responses for RDT (Accept or Reject) and Dots (Correct or Incorrect) were analyzed by fitting logistic multilevel models implemented in the 'lme4' package (Bates, Mächler, Bolker, & Walker, 2015). Maximum likelihood (ML) rather than restricted maximum likelihood (REML) was used in all models, and the Chi-square likelihood ratio difference test was used to assess goodness of fit between nested models (Snijders & Bosker, 2012). Significance values for individual model parameters in CLMMs were calculated using the Wald test, while parameters for logistic multilevel models were estimated using the Welch-Satterthwaite approximation for degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2016). When discussing individually significant predictors in the text, the log odds are converted to the corresponding odds ratio (OR). False discovery rate was used to correct for the multiple planned tests performed across Study 1 (Benajmini & Hochberg, 1995). Significant interactions were investigated separately with post hoc tests using Tukey's HSD correction as implemented in the 'lsmeans' package (Lenth, 2016).

Scenario responses were regressed first on scenario Type, Perspective group, and the Type x Perspective interaction, while controlling for demographic variables. Justice dispositions and sensitivity to reward and punishment were added in two separate blocks. In all models, participant was modeled with a random intercept and a random slope within scenario Type.

For the RDT, acceptance decisions (1 = Accept; 0 = Reject) on individual trials were modeled using mutlilevel maximum likelihood logistic regressions. Two participants were excluded because they failed to provide responses to at least 75% of trials. In the first model, responses were regressed on SelfFairness, OtherFairness, their interaction, and Perspective group, while controlling for age, gender, and SES. Separate terms were included to model random slopes within SelfFairness and OtherFairness in addition to the random intercept for participant. At the second step, dispositional justice sensitivity scores were entered and justice sensitivity scores were allowed to interact with Perspective and the corresponding task level (e.g. JS-Other x OtherFairness, JS-Self x SelfFairness), while controlling for dispositional reward and punishment sensitivity.

Similarly, accuracy (1 = Correct; 0 = Incorrect) for each trial during the Dots task were modeled using hierarchical multi-level maximum likelihood logistic regression. Ten participants were excluded because of poor performance during practice trials (less than 70% correct). Trial accuracy (i.e. correct or incorrect) was regressed on stimulus type (HighReward|LowReward), while controlling for age, gender, SES, accuracy during practice trials, and Perspective group.

Participant was modeled with a random intercept and random slope within each level of stimulus type. Next, justice sensitivity perspectives were entered and allowed to interact with stimulus type. The final step added dispositional reward sensitivity and punishment sensitivity.

Decisions to partake in Freerice (0 = No, 1 = Yes) were modeled using a general linear model with a logit link. Participation was regressed on dispositional other-oriented and self-oriented justice sensitivity, while controlling for demographic variables and Perspective group.

2.3 Results and Discussion

Scenario Ratings. Parameters for the CLMM modeling scenario responses are shown in Table 4. The model with demographic variables (i.e. age, gender, and SES) and task variables (i.e. Type, Perspective, and Type x Perspective) was a better fit than the intercept-only model ($X^2(6) = 27.292$, $p < .001$). Neither justice sensitivity nor approach/avoidance scores improved upon this model (all *ps* > 0.3).

There was a significant effect of scenario Type, with proscriptive transgressions being more likely to elicit moral condemnation (OR = 1.77, $p = .007$). This replicated previous work showing that it is more morally permissible to fail to provide assistance than to engage in harmful or dishonest behavior (Janoff-Bulman et al., 2009). Women were more likely than men to select a harsher rating (OR = 1.56, $p = .029$). Though not a focus of this dissertation, the observed gender effect replicates a previous investigation using similar stimuli (Decety & Yoder, 2015).

The failure to find significant improvement in model fit with the addition of dispositional sensitivity to reward and punishment or justice sensitivity dispositions was unexpected. Reward and punishment sensitivity ostensibly capture individual differences in behavioral approach and avoidance motivations which are the basis for distinguishing between proscriptive and prescriptive morality (Janoff-Bulman et al., 2009). Moreover, the study from which the scenarios were modified identified justice dispositions as increasing explanatory power when modeling moral judgments and as individually significant predictors of those judgments (Decety & Yoder, 2015). Obviously, failing to find a significant increase in model fit does not necessarily mean that there is no effect. Moreover, there are several important differences between Study 1 and the previous study. First, to reduce the time required to complete questionnaires, the short version of the JSI was used, so self-oriented justice sensitivity was assessed with only two responses. Also, Study 1 modeled individual ratings using a multilevel ordinal regression, rather than modeling average permissibility ratings in ordinary least squares multiple regression. Additionally, the previous study used a larger sample (n = 265). Finally, the previous study did not specifically model proscriptive and prescriptive dilemmas separately.

Distributional Fairness. Parameters for the GLMM modeling the impact of justice sensitivity dispositions on decisions in the RDT are shown in Table 5. Including demographic and task variables significantly improved model fit compared to the intercept only model ($X^2(7) = 131.11$, $p < .001$). Adding dispositional scores further improved model fit ($X^2(10) = 29.354$, $p = .006$).

Participants were less likely to accept offers that were unfair to the self (OR = 0.02, $p < .001$). This effect was qualified by a significant OtherFairness x SelfFairness interaction (OR = 22.94, $p < .001$). Post hoc tests revealed SelfUnfair-OtherFair distributions were less likely to be accepted than all other distributions (*p*s < .001; Figure 3). Distributions which were fair to all were more likely to be accepted than those which were unfair to both self and other ($p = .009$). There was no significant difference between SelfFairOtherFair and SelfFairOtherUnfair acceptance rates ($p = .107$) or SelfFairOtherUnfair and UnfairUnfair distributions ($p = .204$).

This suggests that social comparison played an important role because participants were more tolerant of self-unfair distributions if another player in the RDT also received an unfair offer.

There was not a significant main effect of Perspective ($p > .7$), but there was a significant interaction between Perspective and self-oriented justice sensitivity (OR = 0.24, $p = .008$), as well as a three-way interaction between SelfFairness, Perspective, and self-oriented justice sensitivity which became marginally significant after FDR-correction (OR = 0.316, $p = .077$). JS-Self scores were positively associated with acceptance in the Second-Person group, but negatively associated with acceptance in the Third-Person group. To clarify the three-way interaction, I calculated the Kendall rank correlation (τ) between self-oriented sensitivity scores and mean acceptance rates for SelfFair and SelfUnfair distributions within each Perspective group. The *z*-score approximations of the calculated τ values were then directly compared. The positive association between self-oriented justice sensitivity scores and acceptance rates in the Second-Person group was greater for SelfFair than SelfUnfair distributions ($\tau = .12$ and .03, respectively; $z = 3.96$, $p < .001$). There was no such significant difference in the Third-Person group ($p > .4$). This suggests that after evaluating the moral permissibility of actions that would negatively impact oneself, individuals who were dispositionally sensitive to being victimized were motivated to pursue outcomes that benefited the self. This fits neatly with previous articulations of victim sensitivity as a construct. Specifically, it has been argued that higher self-oriented sensitivity translates into a greater likelihood of perceiving injustice (Baumert & Schmitt, 2009) as well as tendency to ruminate on injustice once it has been perceived (Schmitt et al., 1995). In fact, early justice sensitivity research demonstrated a “justice Stroop effect,” where activating concepts of injustice produced slower reaction times to identify justice-related words, but only in individuals with high dispositions justice sensitivity (Baumert, Hangarter, Gollwitzer, & Schmitt, 2007; Hafer & Begue, 2005)

Antisocial Behavior. Parameters for the GLMM modeling the impact of justice sensitivity dispositions on Dots accuracy are shown in Table 6. Including demographic information and task variables, along with accuracy during the practice session, significantly improved model fit over the intercept only model ($X^2(7) = 60.69$, $p < .001$). Adding justice dispositions further increased model fit ($X^2(4) = 15.47$, $p = .016$). Controlling for reward sensitivity and punishment sensitivity scores did result in a significantly better fit ($X^2(2) = 0.10$, $p > .9$).

Unsurprisingly, practice accuracy significantly predicted subsequence accuracy during the task (OR = 4363.19.19, $p = .002$), with individuals who struggled during the practice trials also being less likely to correctly respond during the paid task. Participants were more likely to respond correctly to the HighReward trials (OR = 17.76, $p < .001$), indicating that errors were more likely to be committed specifically in the situation when errors earned participants a higher monetary reward. Women were marginally more likely to provide accurate responses (OR = 1.85, $p = .066$).

There was not a significant interaction between reward type and other-oriented sensitivity (OR = 0.89, $p > .9$). While the failure to find a significant effect is not direct evidence for the nonexistence of that effect, these results are consistent with both the characterization of other-oriented justice sensitivity as a genuine prosocial motivation and previous work linking it to prosocial behavior (Gollwitzer et al., 2005). As expected, there was an interaction between stimulus type and dispositional self-oriented justice sensitivity, though this effect was rendered marginally significant by FDR-correction (OR = 0.35, $p = .082$). Interestingly, the direction of this Type x JS-Self interaction effect was opposite to the hypothesized direction. Individuals with higher dispositional self-oriented sensitivity more likely to respond correctly to LowReward trials (Figure 7). This effect is surprising because self-oriented sensitivity has previously been linked to higher rates of self-reported transgressions (Gollwitzer et al., 2005) and reduced condemnation of moral violations (Decety & Yoder, 2015). Yet self-oriented sensitivity has also been linked to fear of exploitation, which manifests in some contexts as a systematic underestimation of the trustworthiness of others (Gollwitzer et al., 2012, 2015). Thus, it may have been that highly self-oriented individuals were especially wary of the experimental setup (and the presence of the researcher in the room), and may have worked to follow task instructions because they did not believe that they would receive payment for dishonest responses or that dishonest responses would result in some form of punishment.

Prosocial Behavior. Freerice participation was not significantly predicted by other-oriented justice sensitivity, self-oriented justice sensitivity, or Perspective group (see Table 7). Thus, contrary to my hypothesis, there was no evidence to suggest that dispositional other-oriented justice sensitivity was associated with a greater likelihood of engaging in prosocial behavior as measured by participation in Freerice.

Overall, this first study provided partial support for the notion that perspective framing provides cues for sensitive individuals to detect, rather than influencing behaviors directly. Though Perspective group was not an individually significant predictor for moral judgments or unethical behavior during the dots task, there was a marginally significant interaction during the RDT which did improve model fit. Similarly, though dispositional justice sensitivities did not significantly increase model fit for the scenarios task, they did increase fit for Dots responses and (marginally) for RDT responses. Finally, there was a significant interaction between Perspective group and self-oriented justice sensitivity, supporting the notion that dispositional justice sensitivity predisposes individuals to perceive particular sorts of injustice (e.g. self-directed or other-directed), and this renders highly sensitive individuals more susceptible to influences of perspective framing. In other words, Study 1 served as a proof of concept that perspective framing with moral scenarios was able to produce measurable changes in economic decision-making and antisocial behavior as a function of an individual's dispositional justice sensitivity.

2.4 Appendix A: Chapter Two Tables and Figures

Table 1. Study 1 Demographics by Group

	Second-Person (n = 49) Mean (SEM) or %	Third-Person (n = 48) Mean (SEM) or %
Age	35.82 (1.9)	36.4 (2.0)
Percent female	69%	60%
Race/Ethnicity		
White	27%	33%
Black	57%	58%
Hispanic/Latin	8%	4%
Asian	6%	2%
Other or Multiracial	2%	2%

Table 2. Study 1 Dispositional Measures by Group

	Second-Person Mean (SEM)	Third-Person Mean (SEM)	α
JS-Self	1.8 (0.2)	2.1 (0.2)	0.78
JS-Other	2.3 (0.2)	2.1 (0.2)	0.79
RS	5.1 (0.3)	5.4 (0.3)	0.80
PS	4.5 (0.4)	5.0 (0.5)	0.66

Note. JS-Self = Self-oriented justice sensitivity, JS-Other = Other-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. Cronbach's alpha (α) measures internal consistency among individual items, with values above 0.6 considered acceptable. Groups did not significantly differ on any measure.

Table 3. Moral Scenarios

Victim - Second Person Perspective	Observer – Third Person Perspective
Tom is running to catch the bus which only leaves every hour. In front of Tom you are carrying two grocery bags which tear, and their contents spill all over the sidewalk. Besides Tom no one else is around to help. How permissible is it for Tom to catch the bus without helping you?	Tom is running to catch the bus which only leaves every hour. In front of Tom someone is carrying two grocery bags which tear, and their contents spill all over the sidewalk. Besides Tom no one else is around to help. How permissible is it for Tom to catch the bus without helping them?
Mark wants to sell his old car. He knows that the car's engine will need serious repairs pretty soon. You do not notice the problem, but want to buy the car right away and offer to pay in cash. How permissible is it for Mark to sell the car without mentioning the engine problem?	Mark wants to sell his old car. He knows that the car's engine will need serious repairs pretty soon. Someone who does not notice the problem wants to buy the car right away and offers to pay in cash. How permissible is it for Mark to sell the car without mentioning the engine problem?

Table 3, continued

Andrea is taking the final exam for an important class and does not know the answers to several problems. However, Andrea is sitting next to you, the best student in class. Andrea can clearly see your answers. How permissible is it for Andrea to copy your answers?	Andrea is taking the final exam for an important class and does not know the answers to several problems. However, Andrea is sitting next to the best student in class. Andrea can clearly see the other student's answers. How permissible is it for Andrea to copy the other student's answers?
Chris is in a hurry to get to a lunch meeting. Chris gets to a crowded intersection as the crosswalk signal is counting down. Chris can make it by pushing you out of the way. How permissible is it for Chris to push you out of the way?	Chris is in a hurry to get to a lunch meeting. Chris gets to a crowded intersection as the crosswalk signal is counting down. Chris can make it by pushing someone out of the way. How permissible is it for Chris to push them out of the way?

Table 3, continued

Scott is taking the subway train home and is very tired. At the next stop you get on the train with crutches. All of the seats are taken. How permissible is it for Scott to stay in his seat?	Scott is taking the subway train home and is very tired. At the next stop someone with crutches gets on the train. All of the seats are taken. How permissible is it for Scott to stay in his seat?
One of your friends is visiting your house for the first time. There are several pictures of your family members on the walls. Your friend thinks your mother is particularly ugly. How permissible is it for your friend to tell you they think your mother is ugly?	Dana is visiting a friend's house for the first time. There are several pictures of family members on the walls. Dana thinks her friend's mother is particularly ugly. How permissible is it for Dana to say she thinks their mother is ugly?

Table 3, continued

Patricia is your coworker and goes to buy a candy bar from a vending machine. The machine mistakenly gives her two. She knows that you would also like a candy bar, but you didn't ask for one. How permissible is it for Patricia to keep both candy bars?	Patricia goes to buy a candy bar from a vending machine. The machine mistakenly gives her two. She knows that her coworker would also like a candy bar, but didn't ask for one. How permissible is it for Patricia to keep both candy bars?
You and your friend, Justin, want to buy a brand-new computer game. Justin promises to buy each of you a copy of the game. However, when he gets to the store, there is only one game left. How permissible is it for Justin to buy the game for himself?	Justin and his friend want to buy a brand-new computer game. Justin promises to buy each of them a copy of the game. However, when he gets to the store, there is only one game left. How permissible is it for Justin to buy the game for himself?

Table 3, continued

You gave your friend, Sheila, money to buy tickets for a show. However, Sheila forgot. She could say that the show was sold out so that you won't get mad. How permissible is it for Sheila to say the show was sold out?	Sheila's friend gave her money to buy tickets for a show. However, Sheila forgot. She could say that the show was sold out so that her friend won't get mad. How permissible is it for Sheila to say the show was sold out?
Melanie is texting on her phone while walking down the street. She is not paying attention to where she is going and bumps into you. She drops her phone and the screen cracks. How permissible is it for Melanie to yell at you?	Melanie is texting on her phone while walking down the street. She is not paying attention to where she is going and bumps into someone. She drops her phone and the screen cracks. How permissible is it for Melanie to yell at the person?

Table 4. Study 1 Model Parameters for Scenario Responses

Threshold Coefficients	Estimate	SE
1|2	-1.26	0.16
2|3	-0.92	0.16
3|4	-0.53	0.15
4|5	0.22	0.15
5|6	0.50	0.15
6|7	0.98	0.15
Random Effects	**Variance**	**SD**
Participant	0.13	0.36
Type	0.05	0.23
Fixed Effects	**Log Odds**	**SE**
Age	0.13	0.08
Gender (Female)	0.45*	0.17
SES	0.00	0.08
Type (Proscriptive)	0.57**	0.18
Perspective (Third-Person)	0.03	0.18
Type x Perspective	0.23	0.26

Note. ** $p < .01$, *** $p < .001$ (all p values FDR-corrected)

Table 5. Study 1 Model Parameters for Responses in Random Distribution Task

Random Effects	Variance	SD
Participant	1.57	1.25
OtherFairness	9.29	3.05
SelfFairness	2.31	1.52
Fixed Effects	**Log Odds**	**SE**
(Intercept)	4.32***	0.47
Age	0.01	0.24
Gender (Female)	0.86	0.44
SES	0.21	0.21
OtherFairness (Unfair)	-0.43	0.69
SelfFairness (Unfair)	-3.77***	0.47
Perspective (Third-Person)	0.29	0.48
JS-Other	0.25	0.30
JS-Self	0.61	0.39
OtherFairness x SelfFairness	3.13***	0.39
OtherFairness x Perspective	-1.60	0.78
OtherFairness x JS-Other	-0.64	0.50
Perspective x JS-Other	-0.58	0.49
SelfFairness x Perspective	-0.09	0.50
SelfFairness x JS-Self	-0.58	0.39
Perspective x JS-Self	-1.44*	0.54
OtherFairness x Perspective x JS-Other	0.60	0.76

Table 5, continued

SelfFairness x Perspective x JS-Self	1.15†	0.51

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all p values FDR-corrected).

Table 6. Study 1 Model Parameters for Responses in Dots Task

Random Effects	Variance	SD
Participant	5.60	2.37
Type	8.69	2.95
Fixed Effects	**Log Odds**	**SE**
(Intercept)	-6.54**	2.11
Age	0.12	0.14
Gender	0.61†	0.26
SES	0.16	0.13
Practice Accuracy	8.38**	2.38
Perspective (Third-Person)	0.02	0.55
Type (HighReward)	2.88***	0.53
JS-Other	0.38	0.37
JS-Self	0.30	0.37
Perspective x Type	-0.13	0.72
Type x JS-Other	-0.13	0.48
Type x JS-Self	-1.04†	0.47

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all p values FDR-corrected).

Table 7. Study 1 Model Parameters for Freerice Responses

Term	Log Odds	SE
(Intercept)	0.18	0.39
Age	-0.37	0.24
Gender (Female)	0.25	0.49
SES	0.37	0.24
Perspective (Third-Person)	0.25	0.47
JS-Other	0.25	0.27
JS-Self	-0.34	0.28

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity

Figure 1. Stimuli Used in the Random Distribution Task

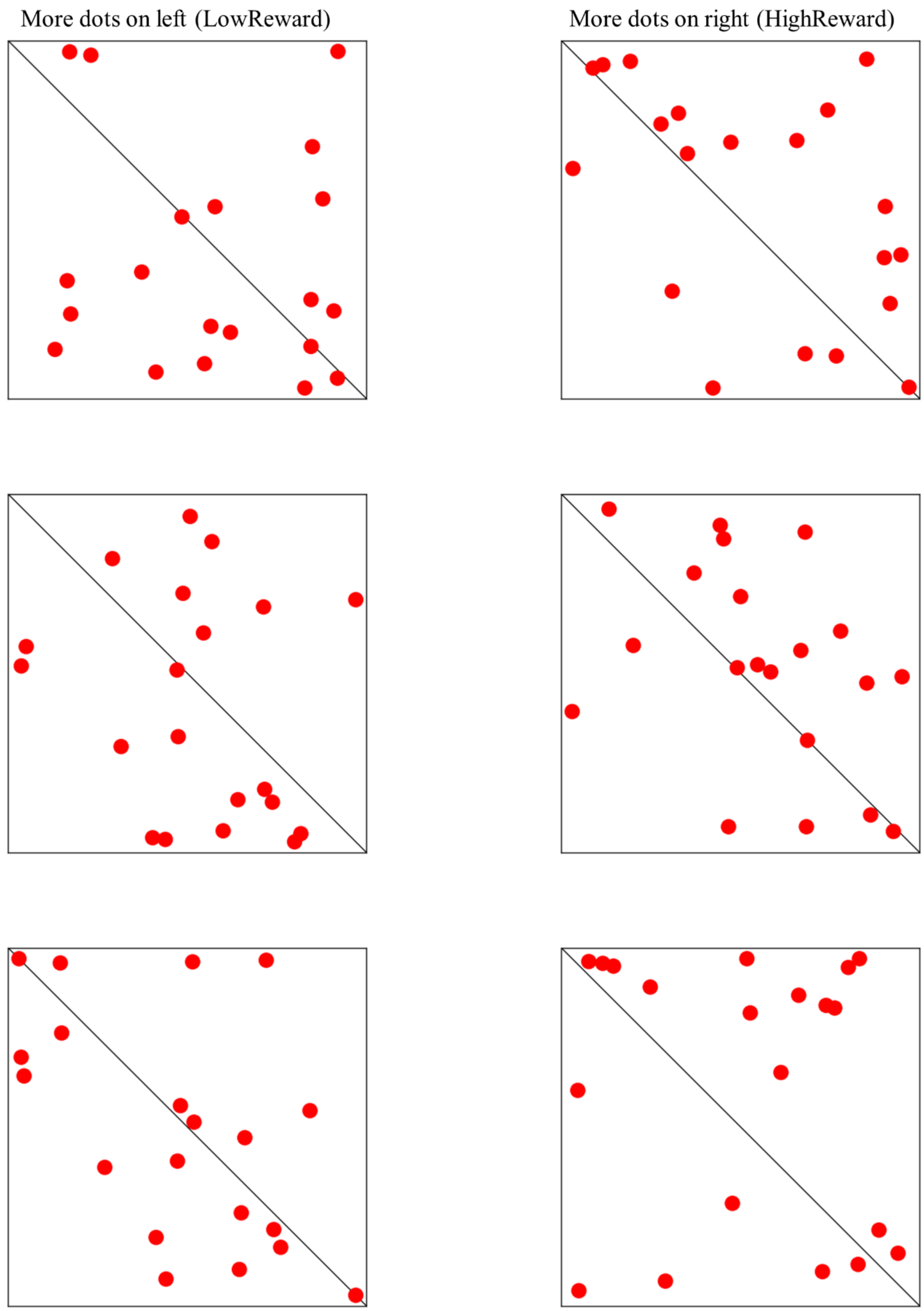


Figure 2. Example Stimuli from the Dots Task

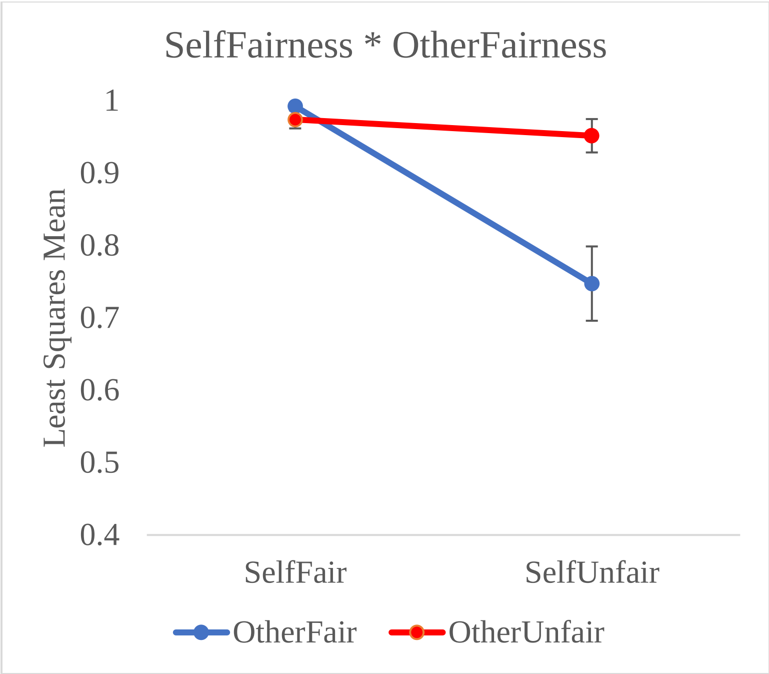


Figure 3. SelfFairness by OtherFairness Interaction in the Distribution Task in Study 1
Least square means for predicted probabilities are plotted with standard error bars.

Chapter 3: Dispositional Influences on Fairness Judgments and Behaviors

3.1 Introduction

Study 1 provided partial support for the importance of perspective framing and justice dispositions for explaining justice-related behaviors, particularly economic decision-making and (resisting) antisocial behavior. The studies in *Chapter 2* sought to replicate and extend these findings using larger online samples while improving on Study 1 in two key ways. Frist, Study 1 utilized the short form of the Justice Sensitivity Inventory, which includes only 8 items. The longer, 40-item version (Schmitt et al., 2010) provides for greater individual variability and potentially offers a more stable and reliable estimate of justice sensitivity dispositions.

Additionally, Study 1 utilized a random distribution task to examine participants' justice-related economic decision-making. However, the moral force behind justice arises, at least in part, because of its social nature (Baumert & Schmitt, 2016; Decety & Yoder, 2017; Sanfey et al., 2003). By informing participants that the distribution allocations had been randomly determined by a computer, participants may have viewed the allocations as divorced from moral rules and social norms about how individuals ought to distribute resources (Buckholtz & Marois, 2012; Gollwitzer et al., 2005; Ruff, Ugazio, & Fehr, 2013). To address this issue, the RDT was modified into a Three-party Distribution Task (TDT). Critically, the TDT added an explicit Proposer, who represented an individual that was given $12 and asked to divide that resource amongst the Proposer, the participant, and a powerless Observer.

3.2. General Methods

3.1.1 Survey administration and participant recruitment

A total of 940 adults completed online questionnaires and tasks across three surveys implemented in the Qualtrics platform (Qualtrics, Provo, UT). Participants were recruited via

Amazon's Mechanical Turk (MTurk; Buhrmester, Kwang, & Gosling, 2011), a crowdsourcing platform where individuals can sign up to perform various tasks in exchange for money. Participants provided informed consent at the start of the survey and were compensated with $1.00 after completion of the survey. MTurk responses were limited to participants in the United States. While the MTurk platform offers researchers an efficient method to collect large amounts of responses quickly, quality control must be exercised when processing responses (see section 3.3.2.4 below).

3.3. Between-Group Effects of Justice Perspectives on Fairness Judgments

3.3.1 Introduction

Study 2 sought to better characterize the impact (if any) of perspective framing and justice sensitivity dispositions on moral judgments, distributional fairness, and antisocial and prosocial behavior. It built on Study 1 by utilizing a larger sample and assessing dispositional justice sensitivity with a more fine-grained measure. Moreover, previous works indicates that empathy also plays an important role in moral decision-making (Decety & Cowell, 2015; Patil & Silani, 2014). However, empathy is a multifaceted construct and includes perspective-taking, motivated concern for others, and affective sharing (Batson, 2009; Decety, 2015; Decety & Jackson, 2004; Derntl et al., 2010; Shamay-Tsoory, 2009). These various aspects of empathy have distinct influences on moral cognition (Decety & Cowell, 2014; Decety & Yoder, 2017), but it is unclear the extent to which they may also influence fairness decisions and antisocial or prosocial behavior. To that end, a measure was added to assess these facets of dispositional empathy in order to better isolate unique contributions of justice sensitivity.

3.3.2 Methods

3.3.2.1 Participants

A total of 300 adults completed an online survey through Amazon's Mechanical Turk (MTurk; Buhrmester, Kwang, & Gosling, 2011) in exchange for $1.00. Participants were restricted to those within the United States. Response analysis (see 3.3.2.4 below) identified 57 participants for removal. Due to technical issues with Qualtrics, an additional 82 participants did not provide complete data, leaving a final sample of 161 participants (see Table 8 for demographic information). Consent was obtained digitally at the beginning of the survey. All materials and procedures were approved by the Institutional Review Board at the University of Chicago.

3.3.2.2 Dispositional Measures

When participants signed up for the study on MTurk they were provided a link to Qualtrics where the questionnaires and tasks were hosted. The SPSRQ-S was again used to assess reward sensitivity and punishment sensitivity as proxies for behavioral approach and avoidance motivation. Unlike in Study 1, justice sensitivity was assessed using the full-length JSI in order to obtain greater variation in justice-related dispositions and a potentially more reliable measure of individual justice dispositions (Appendix G; Schmitt et al., 2010). Participants rated how well statements described them using a 6-point scale (0 = not at all, 5 = exactly), using 10 items for each justice perspective. Scores for victim sensitivity were used to assess self-focused sensitivity (JS-Self), and scores for observer and beneficiary sensitivity were combined to create a single score for other-focused sensitivity (JS-Other). The Interpersonal Reactivity Index (IRI; Appendix H; Davis, 1983) was used to measure three aspects of empathy, each with 8 questions (1=Does not describe me well; 5=Describes me very well). Perspective taking (IRI-PT) captures an individual's ability and propensity to adopt the point of view of another person. Empathic concern (IRI-EC) refers to one's tendency to be motivated to respond to the needs of others. Finally, personal distress (IRI-PD) represents the extent to which another person's negative experiences elicit negative reactions such as anxiety or discomfort. Group means, standard errors, and reliability for each measures are shown in Table 9.

3.3.2.3 Stimuli and Tasks

Moral Framing Scenarios. After completing demographic questions, participants read the same moral scenarios used in Study 1 (see Table 3). As before, participants indicated how morally permissible they thought it would be for one person to behave in a particular situation using a 7-point scale (1 = Completely, 4 = Somewhat, 7 = Not at all). At the start of the survey, participants were randomly assigned to either the Second-Person or Third-Person group, with scenarios written with the corresponding perspective frame (i.e. "you" or "someone").

Distributional Fairness. Stimuli from the RDT used in Study 1 were modified to replace the randomizing computer with a social agent who is making the distributions. In this new Three-party Distribution Task (TDT), one of the other two players was designated as the Proposer and marked with a blue box (Figure 5). The second player was a passive observer. Participants were informed that a decision to "Accept" would result in all three parties keeping the money before them, while a decision to "Reject" would result in no one getting any money. In order to reduce the total amount of time required for participants to complete the online survey and tasks, participants viewed and responded to each hypothetical distribution only once. Participants were allowed to respond at their own pace, and indicated their decision to Accept or Reject by clicking on the appropriate button.

Ethical Behavior. The Dots tasks requires that stimuli be visible for a specified amount of time, and is therefore difficult to reliably employ online were server delays and participants' own computers add unpredictable variability to timing. To address this, studies in *Chapter 3* utilized the Knowledge Misrepresentation Task (KMT), which has previously been used to assess immoral behavior in online studies conducted via MTurk (Reynolds, Dang, Yam, & Leavitt, 2014). The KMT begins by informing participants that they will be answering multiple choice questions about things most eighth graders would have learned in school. For each question, "I don't know" is one of the response options. Two of the questions pose unanswerable questions (e.g. "Who signed the Dente of Bruges?"). The task instructions explicitly ask participants to be honest about their knowledge:

> Below are sample questions from 8th grade textbook test banks on the topics of history, geography and algebra. The questions are rather simplistic, but even so they provide an option for honest but uneducated students to mark "I don't know".
>
> We have hypotheses about the effects of having an "I don't know" option on these questions. It is difficult to test the hypotheses with educated adults, but **in the event that you are unable to solve or answer any of the following questions, please feel free to indicate so by marking "I don't know"**).

Given that participants are explicitly instructed to answer "I don't know" when they cannot identify the correct answer, providing an answer other than "I don't know" is taken to represent a deliberate misrepresentation, i.e. antisocial behavior (mild though it may be). The full set of questions is listed in Appendix I.

Prosocial Behavior. The Freerice website was again used as a measure of prosocial behavior. However, payment on MTurk requires action on the part of the researcher, and it was not feasible to immediately contact MTurk workers immediately following their completion of the survey because of how many individuals would sign up at any given time (e.g. it was not uncommon to receive over 100 respondents in a day). Instead, a second Qualtrics survey was created which explained the purpose of Freerice and loaded the website directly within the survey. A unique survey link was generated for each survey participant, and a follow-up email was sent to each participant through MTurk, several weeks after they had received payment for completing the study. The email stressed that this was entirely optional and that participants could not earn any additional money.

3.3.2.4 Data Analysis

Inattentive participants were identified for exclusion based on four criteria (Curran, 2016). First, total time spend on the survey was used to exclude participants who completed the survey too quickly to have paid full attention (<420 s) or took too much time (>10000 s). Second, participants with suspiciously low variance were identified as those which provided invariant responses on questionnaires that included reverse-scored items. In order to identify participants who provided unusually variable responses, I calculated scores for each measure's odd-numbered and even-numbered items separately. I then compared the two scores, and participants whose odd-even differences were greater than three standard deviations from the mean on any subscale were removed. Finally, an explicit attention check question was included (e.g. "For this question, selection Option 3."). Any participant who failed to correctly answer this question was excluded (all participants in Study 1 answered this question correctly). Any participant whose score on a questionnaire was more than three standard deviations above or below the mean for that questionnaire was marked as an outlier and excluded from analysis. In total, 57 participants were excluded.

Scenario ratings were again modeled using the CLMM approach, regressing individual ratings against predictor variables in an ordinal multilevel model. The first block included demographic and task variables. Justice dispositions were added in the second block, again controlling for reward and punishment sensitivity. Finally, block three added empathy dispositions. In each model, participant was modeled with a random intercept and a random slope within scenario Type.

As in Study 1, I modeled individual decisions during the TDT (1 = Accept, 0 = Reject) using a generalized multilevel model, with separate terms included for random slope within each level of OtherFairness and SelfFairness. Analysis for TDT responses began with the identified model from Study 1 and then added punishment and reward sensitivity dispositions in the second step, and empathy dispositions in the third step.

Ethical behavior on the KMT was assessed by modeling Dishonesty (0 = responded "I don't know" on both target items, 1 = gave an incorrect response to at least one target item) using hierarchical logistic regressions. The first block included the number of incorrect answers to non-target items. In the second block, age, gender, SES, and Perspective group were entered. The third, fourth, and fifth blocks added justice sensitivity, approach/avoidance motivation, or empathy dispositions.

Prosocial behavior was again assessed using a logistic regression to predict whether individuals would respond to the follow-up email (0 = did not respond, 1 = followed link to Freerice survey).

The false discover rate was used to correct for the multiple comparisons in Study 2.

3.3.3 Results and Discussion

Scenario Ratings. Modeling scenario ratings with task and demographic variables was better than then intercept-only model ($X^2(6) = 144.38$, $p < .001$). The block adding dispositional justice sensitivity while controlling for approach/avoidance motivations increased the explanatory power of the CLMM ($X^2(12) = 33.57$, $p = .005$). Empathy scores further improved the model ($X^2(3) = 15.73.60$, $p = .007$). Thus, there was good evidence that justice sensitivity, approach/avoidance motivation, and dispositional empathy are important for predicting participants' permissibility ratings of the scenarios.

Examining individually significant predictors revealed that scenarios describing proscriptive violations were more likely to be rated harshly (OR = 3.84, $p < .001$). Individuals with higher dispositional empathic concern were also more likely to rate scenarios as less permissible (OR = 1.30, $p < .001$). There was also a trend towards dispositional reward sensitivity predicting less harsh ratings, particularly for proscriptive scenarios (OR = 0.87, $p = .100$).

Distributional Fairness. The model from Study 1 was a marginally better fit for the TDT response data that the intercept only model ($X^2(17) = 31.47$, $p = .069$). Adding reward sensitivity and punishment sensitivity scores significantly improved model fit ($X^2(2) = 10.03$, $p = .030$), but empathy dispositions did not ($X^2(3) = 0.93$, $p > .8$).

Offers that were SelfUnfair (allocated more to the Observer than the Participant) were less likely to be accepted than offers which were not (OR = 0.35, $p = .004$). Individuals with higher dispositional punishment sensitivity scores were less likely to accept distributions of any kind (OR = 0.59, $p = .004$). As in study 1, SelfFair-OtherFair offers were more likely to be accepted than SelfUnfair-OtherFair offers ($p < .001$), though the interaction term was rendered marginally significant by FDR-correction (OR = 2.41, $p = .069$; Figure 6).

There was also a marginally significant OtherFairness x JS-Other interaction (OR = 0.42, $p = .094$), wherein individuals with higher dispositional other-oriented justice sensitivity were less likely to accept offers which were unfair for the Observer ($p = 0.004$). This fits with the predictions from Study 1, where other-oriented justice sensitivity was expected to increase rejection of OtherUnfair offers. Justice sensitivity for others reflects an individuals' tendency to react negatively to the injustice of others (Baumert, Rothmund, et al., 2013). Thus, rejections of OtherUnfair offers here may have been motivated by an aversive reaction to these offers.

Since reward and punishment sensitivity dispositions did not significantly improve model fit in Study 1, their inclusion here might suggest that dispositional approach and avoidance are more important for explaining economic decision-making in contexts when allocations are made by a person, rather than a computer. This would fit with previous investigations into two-party UG behavior (Sanfey et al., 2003) and the specific interpretation of rejection decisions as representing altruistic punishment (Baumert, Schlösser, et al., 2014).

Antisocial Behavior. The model for failing to honestly select "I don't know" during the misrepresentation task was significantly improved by including accuracy on the other items of the KMT ($X^2(1) = 14.28$, $p = .001$). Individuals who made more errors on the other multiple choice questions, were more likely to select an incorrect (and ostensibly unethical) choice for either or both of the two target items (OR = 258.01, $p < .001$). This model was not significantly improved by the addition of demographic or task variables, nor any dispositional measure (all $p > .6$). Thus, Study 2 failed to find evidence in favor of a link between justice dispositions or perspective framing on dishonest responding.

Prosocial Behavior. Model parameters for the logistic regression predicting responses to the Freerice follow-up are shown in Table 17. Of the 161 participants, 34 responded (21%). There was a marginally significant effect of age, with older participants being less likely to respond (OR = 0.59, p = .100).

Overall, Study 2 did not find evidence for a direct influence of perspective framing on any of the investigated justice-related behaviors.

3.4 Within-Participant Effects of Justice Perspectives on Fairness Judgments

3.4.1 Introduction

Taken together, the results from Study 1 and Study 2 indicate that when making decisions about monetary distributions, people take into account how the outcome will affect not only themselves, but also others. However, evidence supporting the importance of perspective, whether as a contextual framing effort or as justice dispositions, was mixed. Whereas Study 1 found significant interactions between Perspective group and self-oriented justice sensitivity, Study 2 observed only a marginally significant OtherFairness x JS-Other interaction and no significant main effects of Perspective group. However, in both studies, perspective framing was manipulated between groups, and so these data cannot address the issue of how malleable perspective framing might be. Study 3 sought to address this issue directly by employing a within-participant design.

3.4.2 Methods

3.4.2.1 Participants

440 new adults in the United States participated in a different online survey through Amazon’s Mechanical Turk (MTurk) in exchange for $1.00. I used the same criteria as in Study 2 to identify and remove inattentive responses. The final sample consisted of 370 participants

(see Table 13 for demographic information). As in Study 2, participants signed up for the study on MTurk, then were provided a link to Qualtrics. Signup on MTurk was restricted to within the United States. Informed consent was obtained at the start of the survey. All materials and procedures were approved by the Institutional Review Board at the University of Chicago.

3.4.2.2 Dispositional Measures

Study 3 utilized the same dispositional measures as Study 2: JSI for justice sensitivity, SPSRQ-SF for approach/avoidance motivations, and IRI for empathic dispositions. Descriptive statistics and reliability information are shown in Table 14.

3.4.2.3 Stimuli and Tasks

Moral Scenarios and Distributional Fairness. The structure of the Qualtrics survey was modified to manipulate perspective framing within participants (see Figure 4). In Study 3, participants provided permissibility ratings for two blocks of stimuli. Each block consisted of five scenarios (minimum two scenarios of each type). All scenarios within a block used a consistent Second-Person or Third-Person wording to set up the Perspective frame. Following each scenario framing blocks, participants responded to one block of TDT trials (each containing a mixture of fair and unfair trials for self and other). Block order and the stimuli that were included in each block were counterbalanced.

Antisocial and Prosocial Behavior. Following the second block of TDT trials, ethical behavior was again assessed using the Knowledge Misrepresentation Task (KMT). Rather than waiting to send a follow-up email as in Study 2, prosocial behavior as assessed immediately after the KMT. A screen informed participants that they had finished, and could click a button to advance to receive their confirmation code for payment. Alternatively, they could click a different button and play Freerice. As in the previous studies, it was stressed that participation was entirely optional and that participants could not earn any more money by playing Freerice.

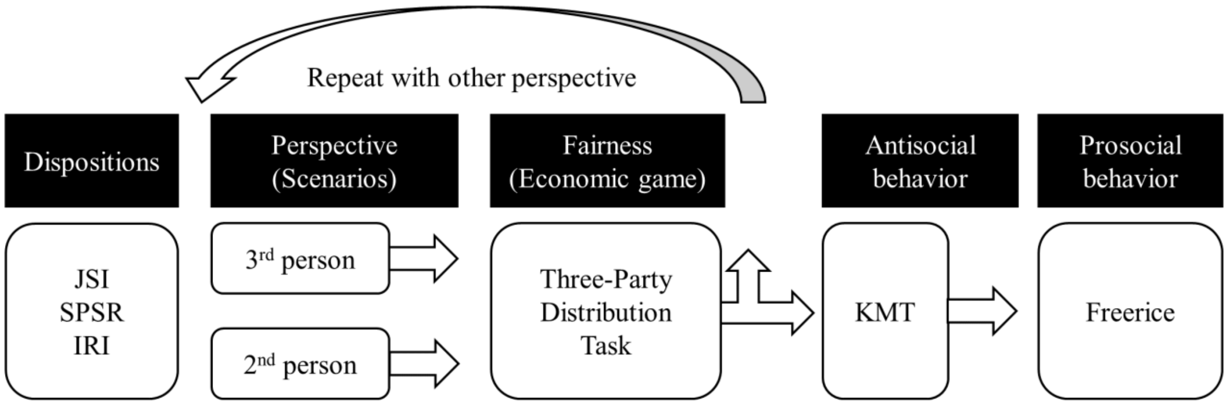


Figure 4. Schematic of Task Order in Study 3.
JSI = Justice sensitivity inventory; SPSR = Sensitivity to punishment/reward; IRI = Interpersonal reactivity index; KMT = Knowledge misrepresentation task. First Perspective was counterbalanced.

3.3.2.4 Data Analysis

Scenario ratings were again modeled in an ordinal multilevel regression using the cumulative link. The nature of a within-participant manipulation of Perspective was unclear, so a stepwise approach to model selection was taken. First, I regressed scenario ratings on Age, gender, SES, Perspective and scenario Type, with Type and Perspective allowed to interact. In successive steps, dispositional justice sensitivity, approach/avoidance motivation, and empathy scores were added, again using the likelihood ratio test to determine if the inclusion of terms significantly improved the explanatory power of the model. Moreover, because Perspective varied within participants, participants were modeled with a random intercept and random slope within both scenario Type and Perspective.

TDT responses (0 = Reject, 1 = Accept) were again modeled using generalized multilevel logistic regressions. As with the scenario ratings, I included a term to model random slopes for each participant within Perspective frame in addition to random slope terms for OtherFairness and SelfFairness. Similarly, justice sensitivity, approach/avoidance motivation, and empathic dispositions were added as separate blocks.

Planned tests were corrected for multiple comparisons using the false discovery rate. Post hoc interrogation of significant interactions was carried out using the Tukey HSD.

3.4.3 Results and Discussion

Scenarios Ratings. Including demographic and task variables significantly improved model fit above the intercept only model ($X^2(6) = 291.94$, $p < .001$). Successive blocks adding justice sensitivity dispositions ($X^2(8) = 28.02$, $p = .003$), reward and punishment sensitivity ($X^2(4) = 13.60$, $p = .043$), and empathic dispositions ($X^2(3) = 35.46$, $p < .001$) each significantly increased model fit. Parameters for the final model are shown in Table 15.

The significance of individual predictors replicated the effects observed in Study 2, with proscriptive violations being more likely to be condemned (OR = 3.84, $p < .001$) and older participants being more likely to use harsher ratings (OR = 1.13, $p = .003$). Also replicating Study 2 was an individually significant effect of dispositional empathic concern, with higher scores increasing the probability that an individual would select a harsher option (OR = 1.30, $p < .001$). There was also a trend towards an interaction between scenario type and reward sensitivity (OR = 0.87, $p = .096$), with higher reward sensitivity associated with greater likelihood of selecting permissible options, especially following Third-Person scenarios.

Distributional Fairness. Model parameters for acceptance of allocations are shown in Table 16. Including demographic and task variables was a marginally better fit than the intercept only model ($X^2(10) = 20.47$, $p = .089$), and adding justice dispositions significantly improved upon that model ($X^2(8) = 31.13$, p = .001). However, including dispositional sensitivity to reward and punishment or empathic scores did not (*ps* > 0.5).

Replicating the previous two studies, there was a main effect of SelfFairness, with participants more likely to reject distributions that were unfair to the self (OR = 0.61, p = .045). This effect was qualified by a significant interaction between SelfFairness and OtherFairness (OR = 2.63, p = .026). There was also a marginally significant OtherFairness x Perspective interaction (OR = 1.86, p = .065), as well as a marginally significant three-way interaction between SelfFairness, OtherFairness, and current Perspective frame (OR = 0.31, p = .084). Post hoc tests did not reveal any significant pairwise differences for the SelfFairness x OtherFairness (*p*s > .1).

Ethical Behavior. As in Study 1 and Study 2, the likelihood of honestly selecting "I don't know" on the two target items was significantly predicted by the number of incorrect responses to non-target items (OR = 86.43, p < .001), but task, demographics, and dispositional measures did not significantly improve the fit of this simple model (all *p*s > 0.4).

Prosocial Behavior. Model parameters for Freerice responses are shown in Table 17. There was a marginally significant effect of other-oriented justice sensitivity (OR = 1.39, p = .093), with higher scores associated with greater likelihood of partaking in Freerice. The inverse effect of self-oriented justice sensitivity scores became non-significant after FDR-correction (OR = 0.75, p = .114).

In summary, Study 3 found some weak evidence to suggest that perspective framing, when manipulated within individuals, may directly impact economic decision-making. However, these effects were only marginally significant, and perspective-framing was not a significant predictor of any other justice-related behavior.

3.5 Resolving Effects of Justice Perspectives on Fairness Judgments: Comparative Equality or Proportional Equity?

3.5.1 Introduction

The goal of Study 4 was to assess whether participants' decisions could be better predicted by modeling "fairness" as comparative equality or proportional equity. Equality refers to situations where the parties receive equal outcomes (Brosnan & Bshary, 2016; Civai, 2013; Decety & Yoder, 2017). Comparative equality here refers to the fact that there are two other individuals (the Proposer and Observer) with which participants might compare themselves. Some previous work suggests that in three-party situations participants do compare their own payoffs to that of others (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999). In contrast, if participants are using what I henceforth call a "proportional equity" standard, they will instead compare their own payoff to what they would receive following the "equal division rule" (Sanfey, 2007). Comparative equality was assumed in Studies 1, 2, and 3 and used as the metric to determine whether a distribution was fair or unfair. For instance, a distribution in which the Proposer kept $6 and offered $2 to the Observer and $4 to the participant (6:2:4) would be classified as OtherUnfair-SelfUnfair, because 2 < 6 and 4 < 6. However, from the standpoint of proportional equity, the participant (and Observer) would not expect to receive more than one third of the initial $12. In other words, the 6:2:4 allocation would be classified as unfair for the other, but *fair* for the self, because 4 >= 4 and 2 < 4. In fact, if proportional equity, rather than comparative equality, is guiding participants' decision-making, then none of the distributions used in the previous studies would be classified as OtherUnfair-SelfUnfair. I addressed this issue in Study 4 by adding a new distribution (10:1:1) which meets the criteria for being unfair to both the Observer and participant, regardless of classification scheme. Additionally, two other distributions were added (6:5:1 and 6:1:5) to examine more extreme instances of unfairness, to provide other distributions which assigned $1 to either the participant or the Observer, and to match other recent investigations in the neural underpinnings of three-party economic decision-making (Alexopoulos et al., 2013, 2012). I then directly compared two models of economic decision-making, one using the comparative equality classification, and the other using the proportional equity classification.

3.5.2 Methods

3.5.2.1 Participants

200 adults took part in Study 4. Following the procedures of Study 3, participants were recruited via MTurk and compensated with $1.00 after completion of the study. Participants provided informed consent at the beginning of the survey. The same response exclusion steps were used as in Study 2: missing data (n = 5), survey duration (n = 7), invariant responding on reverse-scored items (n = 29), odd-even inconsistency (n = 31), and questionnaire outlier (n = 1). This left a final sample of 127 participants (see Table 18 for demographic information). The study was approved by the Institutional Review Board at the University of Chicago, and participants provided informed consent at the beginning of the Qualtrics survey.

3.5.2.2 Dispositional Measures

Study 4 used identical measures to Studies 2 and 3: JSI for justice sensitivity, SPSRQ-SF for approach/avoidance, and IRI for empathy. Descriptive and reliability information is shown in Table 19.

3.5.2.3 Stimuli and Tasks

Moral Framing Scenarios. Study 4 used the same moral scenarios as the rest of *Chapter 3*, with participants rating scenarios in two blocks with each block containing a mixture of prescriptive and proscriptive transgression.

Distributional Fairness. As described above, Study 4 added three new distributions to the Three-party Distribution Task (see Figure 4) in order to adequately model the SelfFairness x OtherFairness interaction when defining fairness as proportional equity.

Antisocial and Prosocial Behavior. Though the purpose of Study 4 was to clarify the fairness metric that best fit the data, antisocial and prosocial tasks where included for completeness and potential comparison to Studies 2 and 3. As in Study 3, the KMT was presented after the second block of TDT trials. An optional Freerice slide at the end provided participants the opportunity to engage in prosocial behavior.

3.5.2.4 Data Analysis

Using Study 3 as a guide, I used an ordinal multilevel model to regress participants' individual ratings on demographics and perspective frame, as well as justice sensitivity dispositions, approach/avoidance motivations, and dispositional empathy scores. Separate terms were included to model a random slope within scenario Type and Perspective frame.

Responses to distributions (0 = Reject, 1 = Accept) were modeled using the GLMM identified in Studies 2 and 3. However, distributions were classified in terms of OtherFairness and SelfFairness in two ways. First, comparative equality was used to classify distributions as before. The second model used the proportional equity classification scheme to label distributions. In both cases, the fairness labels were also used for modeling random slopes within levels of fairness. Model comparison was performed using Akaike Information Criteria (AIC;

Akaike, 1969) because the models were not nested, and thus a likelihood ratio test would have been inappropriate. While the primary goal of Study 4 was to determine whether equity or equality better explained participant's responses during the TDT, for completeness, antisocial behavior during the KMT and prosocial behavior as assessed by Freerice were again modeled with logistic regressions with model parameters shown in Appendix B (Tables 20 and 22).

3.5.3 Results and Discussion

Scenario Ratings. As shown in Table 20, individuals were more likely to rate proscriptive transgressions harshly (OR = 2.64, $p < .001$). There was also a marginally significant effect of dispositional empathic concern (OR = 1.21, $p = .091$) and a marginally significant interaction between reward sensitivity and scenario type (OR = 0.76, $p = .093$).

Distributional Fairness. AIC provides a metric of the quality of various models which have been applied to the same data, and lower numbers indicate a better model. The model using the equity classification had lower AIC that the model using strict equality (1671.70 compared to 1782.30), and was thus selected as the better model of participants' accept/reject decisions in the TDT. For further confirmation, each model was compared to an intercept-only null model with the appropriate error terms. The strict equality model was not significantly better than its null model ($X^2(18) = 23.22$, $p > .4$), but the fairness model did improve over its null model ($X^2(18) = 50.25$, $p < .001$). Parameters from the proportional equity model are shown in Table 21.

In the equity model, only the term for the OtherFairness x SelfFairness interaction was individually significant (OR = 0.09, $p = .002$; Figure 7). Post hoc tests revealed that SelfUnfair-OtherUnfair distributions were less likely to be accepted than distributions that were SelfUnfair-OtherFair ($p < .001$), SelfFair-OtherUnfair ($p = .002$), and marginally less likely to be accepted than SelfFair-OtherFair distributions ($p = .063$). No other pairwise comparison reached significance (*p*s > .2). This result calls into question the social comparison explanation for rejection decisions in Study 1 and Study 2. Moreover, the better fit for the proportional equity model suggests that participants may not have perceived 6:2:4 distributions as unfair to the self. In Study 4, mutually unfair offers were rejected more than any other offer. Especially in light of the difference in AIC, Study 4 provides strong evidence that the equity classification scheme for dilemmas is better at predicting participants' behavior than the strict equality scheme. Moreover, in the TDT, there is little information about the identity of the Proposer and Observer. In such an information impoverished context, it is plausible that participants viewed the Proposer as having some ownership of the initial $12 and thus a justification for keeping more than an "equal" share of $4 as "payment" for their effort (Baumard, André, & Sperber, 2013).

3.6 General Discussion of Surveys

Chapter 3 built on the findings from Study 1 to investigate potential interactions between perspective framing and justice sensitivity dispositions when manipulating perspective between groups (Study 2) or within the same individuals (Studies 3 and 4), while controlling for both approach/avoidance motivation and empathic dispositions. Overall, the findings from Study 1 were partially confirmed. Including measures of dispositional justice sensitivity improved the explanatory fit for models of moral judgment and economic decision-making (Study 2 and 3). There was also evidence for interactions with perspective framing in the context of economic decision-making (Study 3), though the effect was marginally significant.

However, several findings from Study 1 were not replicated. For instance, Study 1 found a marginally significant interaction between unethical behavior and self-oriented justice sensitivity. Here, there was no evidence to suggest that justice sensitivity or perspective framing improved the model's ability to predict which individuals would misrepresent their knowledge on a KMT. In fact, none of the examined variables improved model fit above and beyond participants' accuracy on other items. I do note that there are several important differences between the Dots task in Study 1 and the KMT used in *Chapter 3*, primarily that the Dots task allowed participants to increase their own payoff by behaving unethically, but the KMT offered no specific incentive. Rather, "dishonest" behavior during KMT (Reynolds et al., 2014) could just as easily be explained as careless or random guessing throughout the task. Further, the Dots task was completed while sitting in a room with other individuals present, and social observation has previously been shown to influence behavior (Peterburs et al., 2017).

Importantly, Study 4 demonstrated that proportional equity was better at explaining economic decision-making than comparative equality. This fits with more recent examinations of fairness judgments in economic contexts, where inequality can be justified, and my even be preferred, based on relevant criteria, such as effort or merit (Cappelen et al., 2014; Starmans, Sheskin, & Bloom, 2017). Beginning in childhood, and stabilizing in adolescence, people appear to rely heavily on deservingness when determining whether distributions and allocations are fair (Almås, Cappelen, Sørensen, & Tungodden, 2010; Kanngiesser & Warneken, 2012). With this in mind, *Chapter 4* and *Chapter 5* adopted the proportional equity model when investigating neural systems associated with differential responding to fairness for the self and others.

3.7 Appendix B: Chapter Three Tables and Figures

Table 8. Study 2 Demographics by Group

	Second Person (n = 82) Mean (SEM) or %	Third Person (n = 79) Mean (SEM) or %
Age	36.1 (1.4)	39.6 (1.4)
Percent female	61%	62%
Race/Ethnicity		
White	82%	76%
Black	7%	8%
Hispanic/Latin	4%	3%
Asian	4%	11%
Other or Multiracial	3%	3%

Table 9. Study 2 Dispositional Measures by Group

	Second Person Mean (SEM)	Third Person Mean (SEM)	α
JS-Self	3.8 (0.1)	3.7 (0.1)	0.93
JS-Other	3.9 (0.1)	3.7 (0.1)	0.93
RS	4.5 (0.3)	4.1 (0.3)	0.71
PS	7.2 (0.5)	7.0 (0.5)	0.89
IRI-EC	26.5 (0.6)	26.1 (0.7)	0.90
IRI-PT	24.5 (0.4)	23.7 (0.5)	0.67
IRI-PD	18.7 (0.6)	18.6 (0.7)	0.86

Note. JS-Self = Self-oriented justice sensitivity, JS-Other = Other-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity, IRI-EC = Empathic concern, IRI-PT = Perspective taking, IRI-PD = Personal distress. Cronbach's alpha (α) measures internal consistency among individual items, with values above 0.6 considered acceptable.

Table 10. Study 2 Model Parameters for Scenarios

Threshold Coefficient	Estimate	SE
1|2	-1.73	0.13
2|3	-1.25	0.12
3|4	-0.78	0.12
4|5	0.01	0.12
5|6	0.53	0.12
6|7	1.25	0.12
Random Effects	**Variance**	**SD**
Participant	0.02	0.15
Type	0.03	0.17
Fixed Effects	**Log Odds**	**SE**
Age	0.16**	0.05
Gender (Female)	-0.06	0.11
SES	-0.01	0.05
Type (Proscriptive)	1.18***	0.14
Perspective (Third-Person)	0.20	0.13
JS-Other	0.17	0.12
JS-Self	-0.18	0.11
RS	-0.06	0.07
PS	-0.04	0.07
IRI-EC	0.25***	0.07
IRI-PT	-0.06	0.07

Table 10, continued

IRI-PD	0.09	0.07
Type x Perspective	-0.15	0.19
Type x JS-Other	-0.31	0.16
Perspective x JS-Other	-0.12	0.15
Type x JS-Self	0.37	0.16
Perspective x JS-Self	0.10	0.14
Type x RS	-0.25[†]	0.10
Type x PS	-0.07	0.10
Type x Perspective x JS-Other	0.45	0.23
Type x Perspective x JS-Self	-0.23	0.21

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. [†] $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all *p* values FDR-corrected).

Table 11. Study 2 Model Parameters for Distribution Task

Random Effects	Variance	SD
Participant	1.33	1.15
OtherFairness	2.66	1.63
SelfFairness	0.20	0.45
Fixed Effects	**Log Odds**	**SE**
(Intercept)	1.63***	0.34
Age	0.08	0.16
Gender (Female)	0.34	0.33
SES	0.07	0.15
OtherFairness (Unfair)	-0.30	0.42
SelfFairness (Unfair)	-1.06**	0.30
Perspective (Third)	-0.44	0.34
JS-Other	-0.06	0.26
JS-Self	-0.04	0.27
RS	0.16	0.17
PS	-0.53*	0.18
OtherFairness x SelfFairness	0.88†	0.37
OtherFairness x Perspective	0.15	0.47
Perspective x JS-Other	0.11	0.35
OtherFairness x JS-Other	-0.84†	0.37
SelfFairness x Perspective	0.26	0.34
Perspective x JS-Self	-0.06	0.35

Table 11, continued

SelfFairness x JS-Self	-0.18	0.25
OtherFairness x Perspective x JS-Other	0.16	0.50
SelfFairness x Perspective x JS-Self	0.10	0.34

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all *p* values FDR-corrected).

Table 12. Study 2 Model Parameters for Freerice Responses

Term	Log Odds	SE
(Intercept)	0.18	0.39
Age	-0.37[†]	0.24
Gender (Female)	0.25	0.49
SES	0.37	0.24
Perspective (Third-Person)	0.25	0.47
JS-Other	0.25	0.27
JS-Self	-0.34	0.28

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. [†] $p < .1$ (all p values FDR-corrected).

Table 13. Study 3 Demographic Information

	Sample N = 370 Mean (SD) or %
Age	37.2 (12.4)
Percent female	63%
Race/Ethnicity	
White	82%
Black	8%
Hispanic/Latin	4%
Asian	5%
Other or Multiracial	1%

Table 14. Study 3 Dispositional Measures

	Mean (SD)	α
JS-Self	3.8 (1.1)	0.92
JS-Other	3.9 (1.0	0.94
RS	4.5 (2.3)	0.67
PS	6.8 (4.2)	0.87
IRI-EC	27.1 (5.1)	0.85
IRI-PT	24.4 (3.6)	0.57
IRI-PD	18.1 (5.6)	0.83

Note. JS-Self = Self-oriented justice sensitivity, JS-Other = Other-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity, IRI-EC = Empathic concern, IRI-PT = Perspective taking, IRI-PD = Personal distress. Cronbach's alpha (α) measures internal consistency among individual items, with values above 0.6 considered acceptable.

Table 15. Study 3 Model Parameters for Scenarios

Threshold Coefficient	Estimate	SE
1|2	-1.61	0.15
2|3	-1.05	0.14
3|4	-0.66	0.14
4|5	0.06	0.14
5|6	0.58	0.14
6|7	1.30	0.14
Random Effects	**Variance**	**SD**
Participant	0.05	0.23
Type	0.02	0.13
Perspective	0.00	0.05
Fixed Effects	**Log Odds**	**SE**
Age	0.13**	0.04
Gender (Female)	0.14	0.07
SES	0.00	0.03
Type (Proscriptive)	1.34***	0.19
Perspective (Third-Person)	0.17	0.09
JS-Other	0.33	0.17
JS-Self	-0.24	0.16
RS	-0.02	0.05
PS	-0.06	0.05
IRI-EC	0.26***	0.05

Table 15, continued

IRI-PT	-0.02	0.04
IRI-PD	-0.04	0.04
Type x Perspective	-0.24	0.12
Type x JS-Other	-0.12	0.25
Perspective x JS-Other	-0.11	0.11
Type x JS-Self	0.07	0.24
Perspective x JS-Self	0.07	0.11
Type x RS	-0.14†	0.07
Type x PS	0.05	0.07
Type x Perspective x JS-Other	-0.08	0.16
Type x Perspective x JS-Self	0.08	0.16

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all *p* values FDR-corrected).

Table 16. Study 3 Model Parameters for TDT

Random Effects	Variance	SD
Participant	1.38	1.17
OtherFairness	5.71	2.39
SelfFairness	1.13	1.06
Perspective	0.03	0.18
Fixed Effects	**Log Odds**	**SE**
(Intercept)	1.26***	0.19
Age	0.08	0.09
Gender (Female)	0.16	0.18
SES	0.21[†]	0.09
OtherFairness (Unfair)	-0.35	0.24
SelfFairness (Unfair)	-0.49*	0.19
Perspective (Third-Person)	-0.19	0.17
JS-Other	0.04	0.14
JS-Self	0.14	0.14
OtherFairness x SelfFairness	0.97*	0.34
OtherFairness x Perspective	0.62[†]	0.26
Perspective x JS-Other	-0.21	0.13
OtherFairness x JS-Other	-0.30	0.20
SelfFairness x Perspective	0.39	0.23
Perspective x JS-Self	0.05	0.14
SelfFairness x JS-Self	-0.28	0.15

Table 16, continued

OtherFairness x SelfFairness x Perspective	-1.17†	0.51
OtherFairness x Perspective x JS-Other	-0.25	0.21
SelfFairness x Perpective x JS-Self	-0.24	0.21

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all p values FDR corrected)

Table 17. Study 3 Model Parameters for Freerice Responses

Term	Log Odds	SE
(Intercept)	-0.82	0.39
Age	-0.17	0.12
Gender (Female)	0.23	0.25
SES	-0.02	0.12
Perspective (Third-Person)	-0.14	0.24
JS-Other	0.33[†]	0.15
JS-Self	-0.29	0.14

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. [†] $p < .1$ (all p values FDR corrected)

Table 18. Study 4 Demographic Information

	Sample N = 127 Mean (SEM) or %
Age	36.7 (1.0)
Percent female	64%
Race/Ethnicity	
White	80%
Black	7%
Hispanic/Latin	4%
Asian	6%
Other or Multiracial	3%

Table 19. Study 4 Dispositional Measures

	Mean (SD)	α
JS-Self	3.7 (1.1)	0.92
JS-Other	4.1 (1.0)	0.96
RS	4.5 (2.5)	0.75
PS	7.1 (4.1)	0.86
IRI-EC	26.7 (5.9)	0.90
IRI-PT	24.1 (3.8)	0.60
IRI-PD	18.4 (5.8)	0.86

Note. JS-Self = Self-oriented justice sensitivity, JS-Other = Other-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity, IRI-EC = Empathic concern, IRI-PT = Perspective taking, IRI-PD = Personal distress. Cronbach's alpha (α) measures internal consistency among individual items, with values above 0.6 considered acceptable.

Table 20. Study 4 Model Parameters for Scenario Ratings

Threshold Coefficient	Estimate	SE
1|2	-1.93	0.15
2|3	-1.36	0.14
3|4	-0.92	0.13
4|5	-0.14	0.13
5|6	0.36	0.13
6|7	1.05	0.13
Random Effects	**Variance**	**SD**
Participant	0.15	0.12
Type	0.00	0.02
Perspective	0.01	0.12
Fixed Effects	**Log Odds**	**SE**
Age	0.10	0.06
Gender (Female)	0.18	0.12
SES	0.00	0.05
Type (Proscriptive)	0.97***	0.15
Perspective (Third-Person)	-0.05	0.14
JS-Other	0.05	0.13
JS-Self	-0.21	0.12
RS	-0.09	0.08
PS	0.09	0.09
IRI-EC	0.19[†]	0.08

Table 20, continued

IRI-PT	0.06	0.07
IRI-PD	-0.04	0.07
Type x Perspective	0.09	0.21
Type x JS-Other	-0.08	0.18
Perspective x JS-Other	0.08	0.17
Type x JS-Self	0.07	0.19
Perspective x JS-Self	0.15	0.16
Type x RS	-0.27†	0.11
Type x PS	-0.10	0.12
Type x Perspective x JS-Other	-0.11	0.25
Type x Perspective x JS-Self	0.14	0.25

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all *p* values FDR corrected)

Table 21. Study 4 Model Parameters for Equity Classification of Distribution Task

Random Effects	Variance	SD
Participant	3.85	1.96
OtherFairness	9.97	3.16
SelfFairness	9.60	3.10
Perspective	0.42	0.65
Fixed Effects	**Estimate**	**SE**
(Intercept)	1.53***	0.37
Age	-0.18	0.18
Gender (Female)	-0.19	0.36
SES	0.04	0.18
OtherFairness (Unfair)	0.34	0.46
SelfFairness (Unfair)	0.77	0.47
Perspective (Third-Person)	-0.16	0.27
JS-Other	0.08	0.27
JS-Self	-0.21	0.29
OtherFairness x SelfFairness	-2.43**	0.66
OtherFairness x Perspective	-0.42	0.41
Perspective x JS-Other	-0.20	0.22
OtherFairness x JS-Other	-0.19	0.39
SelfFairness x Perspective	-0.02	0.40
Perspective x JS-Self	0.13	0.23
SelfFairness x JS-Self	-0.76	0.38

Table 21, continued

OtherFairness x SelfFairness x Perspective	-0.25	1.06
OtherFairness x Perspective x JS-Other	-0.07	0.34
SelfFairness x Perpective x JS-Self	-0.11	0.35

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all *p* values FDR corrected)

Table 22. Study 4 Model Parameters for Freerice Responses

Term	Log Odds	SE
(Intercept)	-1.17	0.71
Age	-0.06	0.21
Gender (Female)	-0.26	0.43
SES	0.12	0.20
Perspective (Third-Person)	0.25	0.40
JS-Other	0.24	0.24
JS-Self	-0.13	0.25

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity

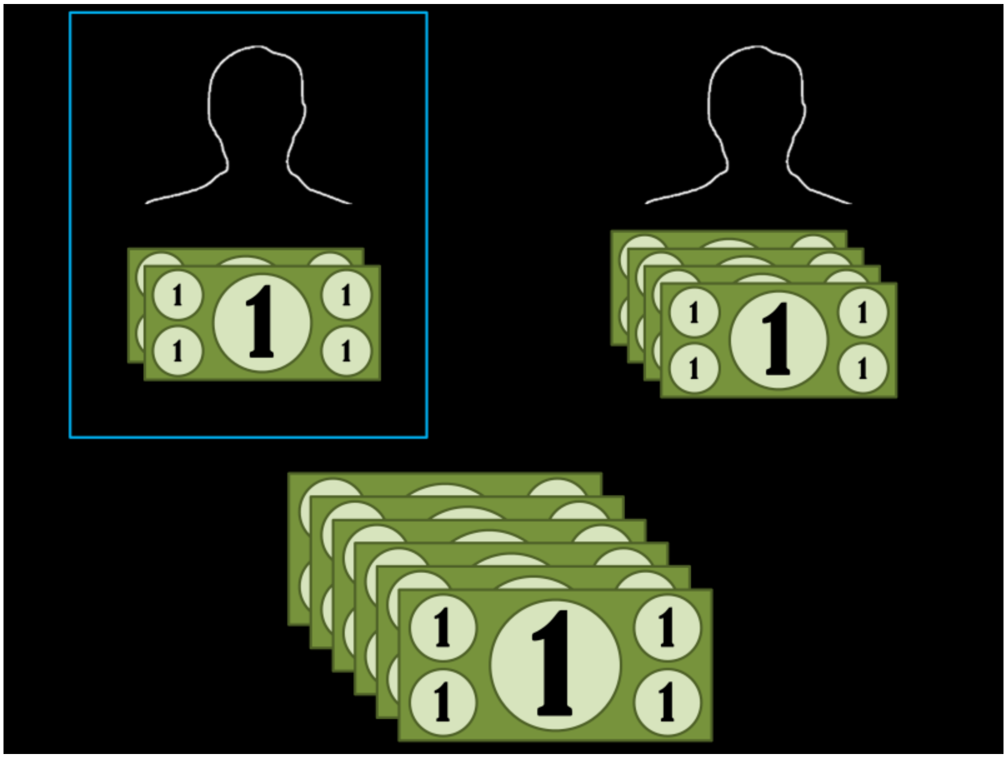


Figure 5. Sample Stimulus for Three-Party Distribution Task

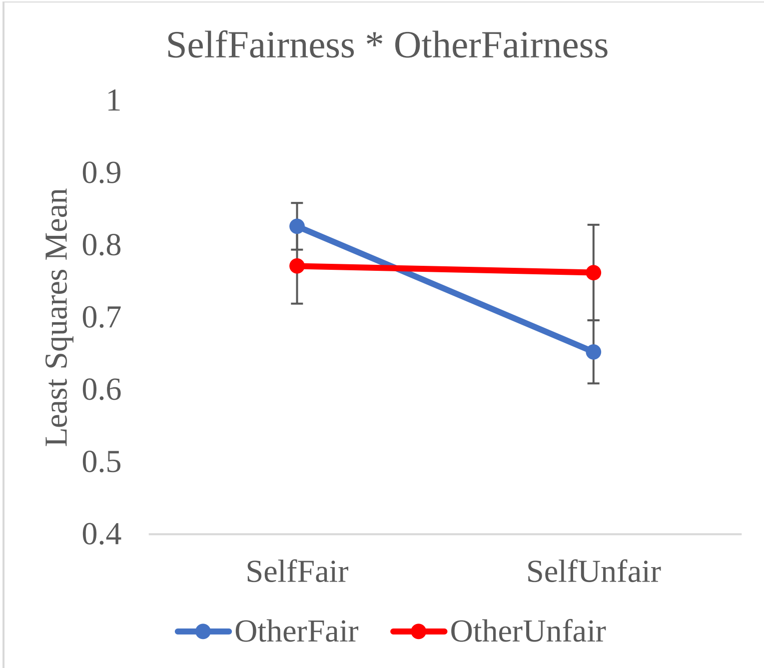


Figure 6. SelfFairness by OtherFairness Interaction from Distribution Task in Study 2
Error bars represent standard error.

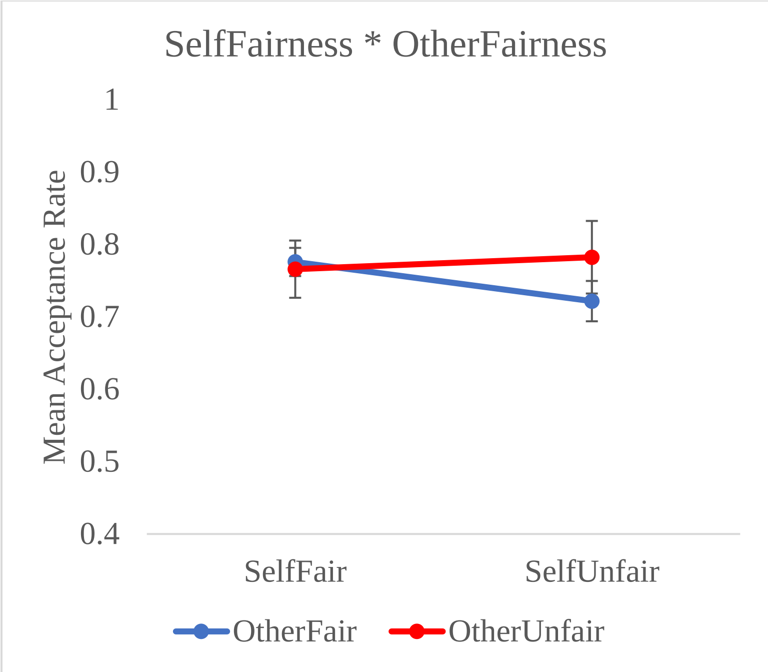


Figure 7. SelfFairness by OtherFairness Interaction from Distribution Task in Study 3
Error bars represent standard error.

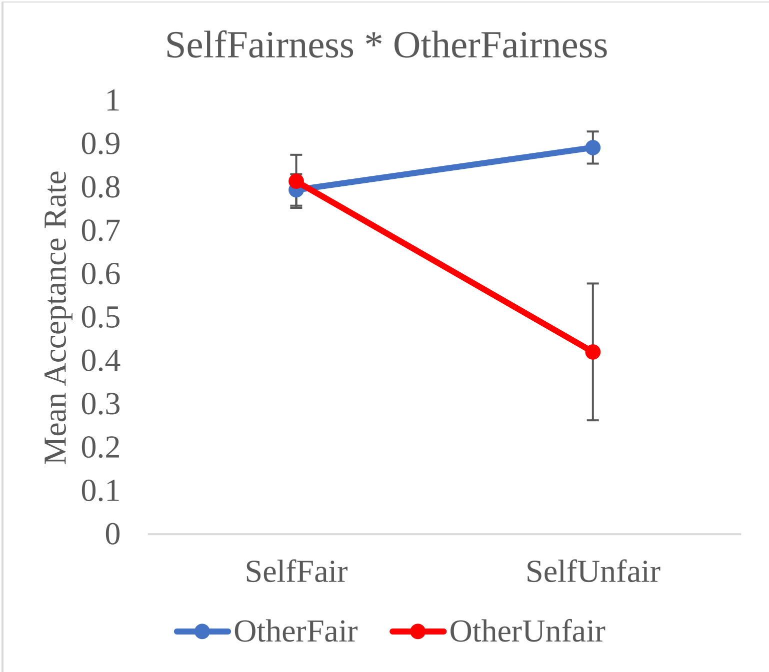


Figure 8. SelfFairness by OtherFairness Interaction from Distribution Task in Study 4
Error bars represent standard error.

Chapter 4: Effects of Justice Sensitivity on Neural Processing of Fairness

4.1 Introduction

Taken together, Studies 1-4 provided consistent evidence in favor of including justice sensitivity dispositions in models of justice-related decision-making. However, in each of these studies, justice sensitivity scores improved model fit, but were not individually significant predictors of behavior after controlling for multiple comparisons. This could suggest that the influence of justice sensitivity on behavior is weak, but could also indicate that dispositional influences are subtle and would thus be difficult to detect using behavioral methods alone. In contrast, fMRI is well-suited to answer questions about the extent to which different aspects of justice sensitivity might relate to common or distinct neural network activity (Mather, Cacioppo, & Kanwisher, 2013). Thus, Study 4 was designed to answer *Question 2* and investigate whether other-oriented and self-oriented justice sensitivity might impact different neural circuitry.

Decades of work in cognitive and affective neuroscience have made great progress in identifying neural networks important for evaluations of fairness. There is now converging evidence from lesions studies (M. Koenigs & Tranel, 2007; Taber-Thomas et al., 2014; Young et al., 2010), studies manipulating neural excitability (Civai, Miniussi, & Rumiati, 2015; Ruff et al., 2013), and functional neuroimaging studies (Decety, Michalska, & Kinzler, 2012; Feng et al., 2015; Gabay et al., 2014; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Moll et al., 2007; Young & Saxe, 2009), that vmPFC, mPFC, ACC, insula, amygdala, pSTS and TPJ each play important and specific roles. Importantly, it now seems clear that decision-making in social contexts relies on the coordination of multiple domain-general systems supporting threat detection, stimulus valuation, response selection, perspective-taking, and mental state understanding (Buckholtz & Marois, 2012; Ruff & Fehr, 2014; Sanfey, 2007; Shenhav &

Greene, 2010). In this emerging view, representations of stimulus and response values and utilities are maintained and updated by circuits involving orbitofrontal cortex, amygdala, and ventral striatum, along with ACC and aINS, and integrated primarily in vmPFC (Balleine & Killcross, 2006; Morrison & Salzman, 2010; Ruff & Fehr, 2014; Wassum & Izquierdo, 2015). Incorporating social information appears to additionally rely on TPJ/pSTS and mPFC, which play critical roles in representing others' mental states, goals, and beliefs, while dlPFC supports goal-directed response selection (Buckholtz & Marois, 2012; Mitchell, Banaji, & Macrae, 2005; Tabibnia et al., 2008; Van Bavel, FeldmanHall, & Mende-Siedlecki, 2015)

Other-oriented justice sensitivity has previously been linked to activity within rTPJ and dmPFC during moral decision-making, as well as increased rTPJ-seeded functional connectivity with ipsilateral parahippocampal gyrus, middle temporal gyrus, and dlPFC (Yoder & Decety, 2014b). This suggests that other-oriented justice sensitivity may primarily influence deliberative processing, rather than modulating saliency or value coding. If this is representative of the mechanisms of justice sensitivity, then higher justice sensitivity should be associated with hemodynamic response within regions important for executive function, such as superior and inferior parietal cortex and dlPFC. However, to my knowledge, this is the only fMRI study to investigate links between hemodynamic activity and justice sensitivity, so it was unclear how well these results would translate to behavioral economics and ethical behavior tasks. For instance, individual differences associated with justice sensitivity might also be associated with changes in the subjective value participants assigned to distributions or their own behaviors.

There is now strong evidence that the nervous system transforms multiple and diverse types of reward into a single "common currency" (K. C. Berridge & Kringelbach, 2015; Levy & Glimcher, 2012; Ruff & Fehr, 2014). This transformation appears to rely on dopaminergic neurons which are capable of integrating both positive and negative information (Schultz, Carelli, & Wightman, 2015). Importantly, response in these neurons scales with individual preferences, mirroring the rank-ordering of food and juice reward (Lak, Stauffer, & Schultz, 2014). In other words, this reward signaling indexes the subjective value assigned to a particular stimulus. Thus, if individuals with high trait other-oriented justice sensitivity find fair outcomes more pleasurable or rewarding, this should manifest as detectable changes in hemodynamic signal within regions known to play a role in valuations, such as striatum and vmPFC, and potentially amygdala and insula (Ruff & Fehr, 2014).

Similarly, victim sensitivity has previously been linked to suspicion of the harmful intentions of others and an enhanced avoidance of negative outcomes for oneself (Gollwitzer et al., 2012, 2015). This suggests that self-oriented justice sensitivity would lead to greater activity in regions linked to avoidance motivation, such as the amygdala and insula (Balleine & Killcross, 2006; Harrison, Hurlemann, & Adolphs, 2015), specifically when evaluating distributions which minimize one’s own payoff. Similarly, self-oriented justice sensitivity involves an expectation that others will behave in exploitive ways (Gollwitzer et al., 2013; Sussenbach & Gollwitzer, 2015). Thus, any observed amygdala activation might also reflect vigilance towards mistreatment (M. Davis & Whalen, 2001; Whalen, 2007). Moreover, distributions that disadvantage the self should be especially aversive (and therefore salient) to individuals with high trait victim sensitivity when self-focused justice principles have been activated following the Second-Person Frame, leading to increased hemodynamic response in dACC and aINS.

Since Study 5 was conducted in the scanner, precise stimulus timing was again available, so the Dots task from Study 1 was used to assess antisocial behavior. Several studies have investigated unethical behavior in the scanner, and they point towards important roles played by the amygdala, striatum, and PFC when engaging in unethical behaviors such as dishonesty and deception (Abe, 2011; Garrett et al., 2016; Sun et al., 2015). In particular, hemodynamic activity in dlPFC and ACC increases during deception in the Guilty Knowledge Task, where participants are instructed to lie about their familiarity with a stimulus (Langleben et al., 2002, 2005). A similar response pattern emerges when participants break a promise in a trust game (Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009) or are asked to lie about their familiarity with both emotionally laden and neutral stimuli (Ito et al., 2011; Lee, Lee, Raine, & Chan, 2010). Finally, dlPFC response increases when participants anticipate telling a lie, pointing towards its preparatory role (Ito et al., 2012). Thus, dlPFC activity has been argued to reflect executive functioning which is required to reduce the emotional conflict created by behaving dishonestly (Greene & Paxton, 2009). In other words, telling the truth is the "natural" baseline state, and going against that requires extra effort. However, one recent study suggests that repeated lying gradually degrades the emotional impact of being dishonest, with self-serving dishonesty increasing over time, coincident with reduced response in the amygdala (Garrett et al., 2016)

Here, where stimuli are distinguished not only by their physical properties, but also by their actual monetary payoffs, neural regions involved in stimulus valuation were also obvious candidates for potential influences of justice sensitivity. In particular, self-oriented justice sensitivity was expected to predict enhanced hemodynamic response for HighReward compared to LowReward trials within the ventral striatum and vmPFC. Conversely, as it ostensibly indexes general prosocial motivation, individuals high in other-oriented justice sensitivity were expected to find the prospect of dishonestly responding to LowReward trials less tempting, and so JS-

Other scores were expected to be negatively associated with activity within regions involved in response inhibition during LowReward trials, especially the extreme ventral extent of IFG (Aron, Robbins, & Poldrack, 2014; Sebastian et al., 2016). Moreover, individuals with higher self-oriented sensitivity were expected to provide more dishonest responses and show reduced activity in dlPFC and amygdala during LowReward trials.

4.2 Methods

4.2.1 Participants

A total of 32 adults participated in the fMRI study and were compensated with a combination of cash and gift cards (see Table 23 for demographic information). Participants were each paid $20, plus whatever they earned during the Dots task (mean $5.11). Exclusion criteria included MRI contraindications (e.g. metal implants), and self-reported history of psychoactive medication or head injury resulting in unconsciousness lasting more than 10 minutes. Study materials and procedures were approved by the Institutional Review Board at the University of Chicago. All participants provided informed written consent.

4.2.2 Dispositional Measures

After phone screening for MRI eligibility, and approximately one week prior to scanning, participants completed dispositional measures online, again using the Qualtrics platform. As before, self-oriented and other-oriented justice sensitivity were assessed using the full-length JSI, reward and punishment sensitivity were indexed using the SPSRQ-SF, and empathic concern, perspective-taking, and personal distress were measured using the IRI. Descriptive statistics and reliability information are shown in Table 24.

4.2.3 Stimuli and Tasks

Moral Framing Scenarios. As in Studies 3 and 4, prior to distribution and antisocial behavior tasks, participants first rated the moral permissibility of scenarios, written in either second-person or third-person. Responses were given using an MR-compatible button box to select one of seven options (1 = Completely, 4 = Somewhat, 7 = Not at all). Participants were allowed to respond at their own pace, and functional images were not acquired during responses. After the scenarios, participants completed one run of the TDT, followed by one run of the Dots task, all while undergoing fMRI. Next, a structural image was acquired, during which time participants were allowed to rate the second set of scenarios, before completing the final TDT and Dots scanning runs.

Distributional Fairness. The TDT was adapted for use with fMRI. In keeping with the model identified in Study 4 and previous neuroimaging investigations for three-party fairness (Alexopoulos et al., 2012; Ma & Hu, 2015), a "Fair" offer was defined as at least one third of the initial endowment ($4 or more), while an offer of less than 15% (i.e. $1) was labeled "Unfair." Four exemplar distributions were chosen (4:4:4, 6:5:1, 6:1:5, and 10:1:1; see Appendix J) to fill the 2 Self (Unfair|Fair) x 2 Other (Unfair|Fair) design. Importantly, these distributions also increased the severity of SelfUnfair offers compared to Studies 1, 2, and 3. This was done to increase the strength of response. In the scanner, distributions were presented for 3.5 seconds. Participants were instructed to indicate their decision to "Accept" or "Reject" the distribution by pressing the appropriate button as soon as they reached a decision. In each run, stimuli order was randomized. Participants always played the role of the Responder, but the Proposer could be on the left or the right of the display. Each stimulus type was presented 16 times (half with left Proposer, half with right Proposer), resulting in a total of 64 trials per run. Trials were separated by a jittered fixation cross (*M* = 4 s, *SD* = 1.1 s), and participants were given a 10 s break in the middle of each run.

Antisocial Behavior. Unlike online surveys, stimulus presentation during MRI acquisitions allows for fine-grained control of the timing. Thus, the Dots task from Study 1 was adapted for fMRI to assess antisocial behavior. Two new sets of stimuli were generated. Prior to scanning, participants completed 50 practice trials while seated at a computer. All participants demonstrated at least 92% accuracy (no more than 4 out of 50 incorrect responses) during the practice trials, demonstrating ability to perform the task. As in Study 1, participants were not informed of the payment schedule until after completing the practice trials. The payment amounts were adjusted (LowReward = 0.5 cents, HighReward = 5 cents) to match those of previous investigations using the task (Gino & Ariely, 2012; Mazar & Zhong, 2010).

Inside the scanner, participants completed a total of two runs of the Dots Task while undergoing fMRI. Each run contained 100 stimuli (60 Lowearn and 40 Highearn). The timing was similar to Study 1, with each array of dots presented for one second, followed by a response screen which remained visible for two seconds. Participants indicated the side with more dots using an MR-compatible button box. OptSeq2 was used to select stimulus onset times in order to maximize design efficiency while keeping the total task under ten minutes (interstimulus interval *M* = 3.2 s, *SD* = 3.1 s). Each run contained a different set of images, such that participants never saw the same dots display twice across all practice or functional runs. The stimulus-reward mapping (e.g. whether HighEarn was “right” or “left”) and stimulus set used were counter-balanced between participants.

Prosocial Behavior. I had originally planned to present participants with an opportunity to play Freerice immediately after scanning. However, because of a combination of factors, including scheduling, available space, and construction, this was not feasible. Instead, as in Study 2, a follow-up email was sent to participants providing instructions and a link to Freerice. Again, it was stressed that participation was optional and no additional compensation was available.

4.2.4 MRI Acquisition and Processing

Participants were scanned with a 3T Phillips Achieva scanner located at the University of Chicago's MRI Research Center. High-resolution T1-weighted anatomical scans were acquired using a 3D MP-RAGE sequence (repetition time = 8ms; echo time = 7 ms; voxel size = 0.9 x 0.9 x 0.9 mm^3; matrix = 256 x 256). Function images were acquired along the transverse plane oriented to the AC-PC line using a single-shot EPI sequence (voxel size = 3.5 x 3.5 x 3.5 mm^3; skip gap = .5mm; flip angle = 77^o; matrix size = 64 x 64; echo time = 26 ms; repetition time = 2 s).

Functional images were first realigned and high-pass filtered (cutoff = 128 s). Next, the mean realigned image was coregistered to the participant's structural scan. Structural scans were then segmented into five tissue types, and tissue types were separately normalized to the SPM12 structural template. The warping parameters obtained from this process were then applied to the functional images before smoothing using a 7 mm full width half maximum (FWHM) Gaussian smoothing kernel.

4.2.5 Data Analysis

4.2.5.1 Analysis of Behavior Data

The moral scenarios were used here primarily to manipulate participants' perspective prior to scanning runs. Moreover, the sample size is greatly reduced compared to *Chapters 2-3*, so attempts to fit a multilevel cumulative link model to these data failed to converge. However, for completeness, scenario ratings were treated as interval scaled and averaged for each participant within each Type x Perspective cell so they could be compared to the preceding results. Averaged participant ratings were interrogated using hierarchical linear multilevel models, using a random intercept for each participant. In a stepwise fashion, demographic and task variables were entered first, followed by justice dispositions.

Responses during TDT were assessed using the model identified in Studies 3 and 4. Briefly, TDT responses (0 = Reject, 1 = Accept) were regressed on justice sensitivity dispositions, as well as their interactions with OtherFairness, SelfFairness, and current Perspective. Random slopes were included for OtherFairness, SelfFairness, and Perspective.

Similarly, Dots responses (0 = Incorrect, 1 = Correct) were modeled using a generalized multilevel linear regression with a logit link. The model identified in Study 1 was expanded to include dispositional reward and punishment sensitivity, in order to isolate any independent effects of justice sensitivity dispositions. Terms were included for random slopes within each level of Type and Perspective.

Response to the Freerice follow-up email (0 = Did not respond, 1 = Did respond) was examined with a logistic regression.

4.2.5.2 Neuroimaging Analysis

Analysis of functional data was conducted with general linear models (GLMs). At the first-level, stimulus onsets and durations were modeled with a boxcar function and convolved with the canonical hemodynamic response function (HRF) in SPM12. TDT trials were modeling beginning at the onset of the distributions slide and ended with the offset of the slide (duration = 3.5 s). Dots trials were modeled beginning at the onset of the dots display and ending with the offset of the feedback slide (duration = 3 s). Importantly, only trials to which participants responded were included in analysis. Movement parameters were modeled as nuisance regressors. Using the ArtRepair toolbox (Mazaika, Hoeft, Glover, & Reiss, 2009), functional images with more than 0.5 mm/TR were interpolated and a separate regressor for every such image was added to the first-level design matrix in order to "deweight" these images. If any single functional run required repairs of more than 20% of the volumes, that participant was not included in analysis. This excluded two participants from each task (one participant required excessive correction for both tasks).

A threshold of $P < .005$ with a cluster extent of 10 was selected for whole-brain statistical maps because this level has been suggested as the optimal balance between Type I and Type II errors in fMRI studies (Lieberman & Cunningham, 2009). This threshold was adjusted based on the number of planned second-level contrasts: $P < .001$ for five contrasts in TDT (OtherUnfair – OtherFair, SelfUnfair – SelfFair, OtherFairness x SelfFairness, Perspective x OtherFairness, and Perspective x SelfFairness) and $P < .0025$ for two contrasts in Dots (HighReward – LowReward and Perspective x Type). At the second level, self-oriented and other-oriented justice sensitivity scores were entered as covariates, with age and dispositional reward and punishment sensitivity included as covariates of no interest.

4.3 Results and Discussion

4.3.1. Behavioral Responses

Scenario Ratings. Modeling participant responses using the linear multi-level model with demographic and task variables trended towards a better fit than the intercept-only model ($X^2(5) = 14.22$, $p = .070$; Table 25). Adding justice dispositions did not significantly improve model fit ($p > .7$). Participants were marginally more likely to rate proscriptive violations as worse than prescriptive violations ($\beta = 0.54$, $p = .078$).

Distributional Fairness. Table 26 shows the model parameters for responses during the TDT. There were significant main effects of OtherFairness (OR = 0.00, $p < .001$) and SelfFairness (OR = 0.01, $p < .001$), which were qualified by a significant OtherFairness x SelfFairness interaction (OR = 49.63, $p < .001$; Figure 9). Offers that were fair to both were more likely to be accepted than any other offer (*p*s < .001). SelfUnfair-OtherFair offers were more likely to be accepted than both SelfFair-OtherUnfair ($p = .001$) and SelfUnfair-OtherUnfair distributions ($p < .001$). There was not a significant difference between SelfFair-OtherUnfair and SelfUnfair-OtherUnfair distribution acceptance rates ($p > .2$).

Participants were more likely to accept distributions following Third-Person scenarios, though this effect became marginally significant after multiple comparison correction (OR = 6.55, $p = .070$). However, there was a significant three-way interaction between OtherFairness, Perspective, and other-oriented justice sensitivity (OR = 0.07, $p < .001$), with higher other-oriented justice sensitivity more strongly associated with differences in rejection rates between OtherFair and OtherUnfair offers in the Third-Person Frame than the Second-Person Frame ($p = .049$). This supports the notion that other-oriented justice sensitivity predisposes individuals to detect injustice directed at others (Baumert & Schmitt, 2009). Moreover, it extends this work and shows that adopting an other-focused perspective changes how individuals make decisions in subsequent situations where fairness concerns for the self and other conflict.

Antisocial Behavior. Overall, participants were quite honest during the Dots task (M = 183.6 (92%) correct, SD = 29.6). As shown in Table 27, there was a marginally significant main effect of punishment sensitivity (OR = 0.25, $p = .070$), which was qualified by a significant Type x punishment sensitivity interaction (OR = 7.41, $p = .020$). Dispositional punishment sensitivity increased accuracy on HighReward trials (Kendall's $\tau = .33$), but decreased accuracy on

LowReward trials (τ = -.26, z = 13.66, p < .001). Previous investigations in reward processing highlight the role subjective reframing can play (Pessiglione & Delgado, 2015), where a reduction in reward outcome can be experienced as an aversive punishment. Thus, individuals with higher punishment sensitivity may have altered their behavior to avoid the less rewarding (i.e. punishing) payment outcome.

Prosocial Behavior. None of the predictors in the logistic regression modeling responses to the follow-up email reached significance (see Table 28 for full model parameters).

4.3.2. Hemodynamic Responses

Distributional Fairness. Significant activations in the OtherUnfair > OtherFair contrast included clusters in left angular gyrus, and precuneus (Table 29), two regions which overlap with a recent meta-analysis of fMRI studies using classic two-party Ultimatum Games (Feng et al., 2015). Interestingly, there was no significant activation within the core nodes of the salience network or reward processing regions. There were also no significant clusters identified in the OtherFair > OtherUnfair contrast. Other-oriented and self-oriented justice sensitivity scores were associated with overlapping clusters in the posterior portion of left striatum (see Figure 10). Activity in this cluster was greater when evaluating offers which were unfair to the Observer, and was enhanced by other-oriented justice sensitivity, but negatively associated with self-oriented justice sensitivity. While the striatum is often implicated in processing positive outcomes, dopaminergic signaling, including subtypes of neurons within the striatum, respond to reward, punishment, or both (Matsumoto & Hikosaka, 2009; Pessiglione & Delgado, 2015). This cluster overlaps with an area of posterior striatum which is more responsive to monetary losses than gains (Seymour, Daw, Dayan, Singer, & Dolan, 2007). Therefore, I take this as evidence that distributions which assigned $1 to the Observer were assigned more negative valence by individuals with higher other-oriented justice sensitivity, but less negative valence by individuals with higher self-oriented justice sensitivity. To my knowledge, this is the first direct evidence that the proposed prosocial and antisocial consequences of other-oriented and self-oriented justice sensitivity may be associated with patterns of striatal response.

When examining the main effect of SelfFairness, multiple regions showed significantly greater response during SelfFair compared SelfUnfair distributions (Table 30). Right ventral striatum showed greater activity for SelfFair offers, consistent with its role in encoding rewarding outcomes (O'Doherty, 2004; O'Doherty, Deichmann, Critchley, & Dolan, 2002). The hedonic encoding interpretation is also consistent with observed activity in right parahippocampal gyrus, centered in hippocampus, but extending into right amygdala (Simon et al., 2010). While there was significant cingulate activity, the cluster was in posterior midcingulate, an area functionally linked with the centromedial complex of the amygdala and implicated in response selection of rewarding behavior (Bzdok, Laird, Zilles, Fox, & Eickhoff, 2013). There were no significant clusters identified in the SelfUnfair > SelfFair contrast. Moreover, no regions showed significant associations with justice sensitivity dispositions in either SelfFairness contrast.

Table 31 shows clusters with a significant SelfFairness x OtherFairness interaction effect. Regions identified by this contrast overlapped with the main nodes of default mode network, including bilateral angular gyrus, precuneus, and medial prefrontal cortex (Buckner, Andrews-Hanna, & Schacter, 2008; Fox et al., 2005; Raichle, 2015). As shown in Figure 11, each of these regions showed a similar pattern, with greater hemodynamic activity for OtherUnfair than OtherFair during SelfFair trials, but similar OtherUnfair-OtherFair differentiation during

SelfUnfair trials. These regions play an important role in social cognition and emotional processing (Schilbach et al., 2012), as well as theory of mind and moral cognition (Bzdok et al., 2012; Greene et al., 2001; Saxe & Kanwisher, 2003). Given these diverse functions, activity within this network is complicated to interpret (Hutzler, 2014), yet is consistent with an interpretation in which individuals consider their own payoffs, social norms, and potential reputational consequences when deciding whether to accept a distribution which benefits them at the expense of the Observer.

Self-oriented justice sensitivity significantly predicted reduced interaction effects ($\text{SelfUnfair}_{\text{OtherUnfair-OtherFair}} – \text{SelfFair}_{\text{OtherUnfair-OtherFair}}$) in a host of regions (Table 31). This included a large cluster in the brainstem, which extended through the left midbrain into striatum, anterior insula, and amygdala. This interaction effect was also evident in bilateral ventral striatum (Figure 12) and arose because self-oriented sensitivity predicted greater striatal responses to OtherUnfair compared to OtherFair allocations during SelfFair trials, but a reversal for SelfUnfair trials (*p*s < .001). One interpretation of this finding is that individuals with higher self-oriented sensitivity assigned more value to others' unfair outcomes if they themselves simultaneously received an unfair allocation. In other contexts, striatum is critical for representing prediction errors (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Schönberg, Daw, Joel, & O'Doherty, 2007), and striatal prediction errors are drivers for avoidance learning, with difference in striatal activity predicting whether individuals will demonstrate appetitive or aversive learning style (Eldar, Hauser, Dayan, & Dolan, 2016). Thus, the striatal modulation observed here could reflect processing in the context of expectations about the Proposer, and suggest that highly victim-sensitive individuals interpreted SelfUnfair-OtherUnfair offers as evidence that they had not be singled out (K. C. Berridge & Kringelbach, 2015).

Moreover, these data suggest that self-oriented justice sensitivity may also be associated with changes in physiological states. The large midbrain cluster extended into the locus coeruleus (LC), and there was a separate significant cluster with a peak in the supplementary motor area (SMA; Figure 12). The LC is the primary source of norepinephrine (NE) for the forebrain and the LC-NE system has innervations throughout the central nervous system (Schwarz & Luo, 2015). LC discharges are important for a broad range of state changes, including waking, and the LC-NE system plays an important role in modulating cortical and subcortical processing of salient information (C. W. Berridge, 2008; C. W. Berridge & Waterhouse, 2003). In fact, the recently proposed glutamate amplifies noradrenergic effects (GANE) model argues that the ascending LC-NE enhances the processing of salient information by selectively improving excitatory glutamatergic signalling for high priority representations while inhibiting less active representations (Mather, Clewett, Sakaki, & Harley, 2016). In the GANE model, general physiological state changes in response to endogenous or exogenous stimuli, whether they are appetitive (e.g. eating ice cream) or aversive (e.g. remembering an embarrassing moment), lead to increases in NE, which enhances information processing selectivity. In the sensitivity to mean intentions (SeMI) model of victim sensitivity, individual differences in self-oriented justice sensitivity are argued to reflect not only an increased tendency to react negatively to exploitation, but also an expectation that others harbor malicious intentions (Gollwitzer & Rothmund, 2011; Gollwitzer et al., 2013). In light of the GAIN model, this association with LC could represent a neural mechanism by which the strength of negative reactions to perceived exploitation could sharpen information processing specifically for injustice cues.

Finally, in addition to its role in the salience network (Shackman et al., 2011), the SMA/MCC is implicated in the central autonomic network (CAN), a network of brain regions important in interoception and the regulation of autonomic functioning (Benarroch, 1993). During tasks which are both cognitively difficult or emotionally evocative, this region correlates with sympathetic, but not parasympathetic activity as measured by electrodermal activity (EDA) and high-frequency heart rate variability (HF-HRV), respectively (Beissner, Meissner, Bär, & Napadow, 2013). Conversely, caudate activity is reliability associated with parasympathetic response. Parasympathetic activity, as assessed with changes in HF-HRV, has also been argued to be an important part of a generalized threat response (Thayer & Lane, 2009). In other words, SMA/MCC response here could reflect a tendency for individuals with higher self-oriented justice sensitivity to perceive unfairness as threatening.

Regions showing greater hemodynamic response for OtherUnfair > OtherFair following the Third-Person frame are shown in Table 32. One of these clusters was in anterior midcingulate cortex (aMCC). This region overlaps with the posterior extent of a dACC, an area activated by cognitive control, pain, and negative affect (Shackman et al., 2011), and one of the core nodes of salience network discussed in *Chapter 1* and the introduction to *Chapter 4*. The increase in this region for OtherUnfair – OtherFair was greater after the Third-Person frame, suggesting that distributions that negatively affected the Observer were perceived as more salient after the Third-Person Frame compared to the Second-Person Frame. Such an effect suggests that the perspective framing manipulation was successful at biasing subsequent information processing.

Antisocial Behavior. Table 33 lists the significant clusters which showed greater activity for HighReward compared to LowReward Dots stimuli. Greater hemodynamic response was observed in left precentral and postcentral gyrus, likely representing the contralateral motor preparation and execution of the button press. Moreover, as expected, there was a significant cluster in the left striatum, suggesting that individuals did encode the higher reward value the HighReward stimulus. There were no regions showing significant activity in the LowReward > HighReward contrast.

Dispositional justice sensitivity scores demonstrated largely divergent influences on hemodynamic activity for the HighReward – LowReward contrast. Other-oriented justice sensitivity predicted increased signal change in bilateral supramarginal/TPJ and IFG (Figure 13), nodes of the stimulus-driven ventral attention network (Corbetta & Shulman, 2002; Fox, Corbetta, Synder, Vincent, & Raichle, 2006; Vossel, Geng, & Fink, 2014). Given the ventral attention network's role in reorienting to infrequent or unexpected behaviorally relevant stimuli (Vossel et al., 2014), this activation might suggest that individuals with higher dispositional concern for the injustice of others were more sensitive to the stimulus presentation schedule (only 40% of trials were HighReward). Moreover, the activity in IFG is consistent with my hypothesis that individuals with higher other-oriented justice sensitivity would be less motivated to respond dishonestly, and therefore require less prefrontal inhibition (Sebastian et al., 2016). JS-Other scores also predicted enhanced neural response in bilateral striatum (Figure 14). Increased striatal response likely reflects a positive prediction error in response to receiving a rarer HighReward trial, and could indicate that other-oriented sensitivity was associated with increased subjective value of these trials (Lak et al., 2014). This would fit with the hypothesized link between prosocial motivation and assigning hedonic value to trials which allowed higher, guilt-free rewards.

Conversely, self-oriented justice sensitivity scores were associated with widespread reductions in hemodynamic response for HighReward > LowReward. This included integrative hubs in ACC and left TPJ (Figure 15), as well as bilateral insula and parahippocampal gyri (Figure 16). This effect also manifested in midcingulate cortex. Unlike dACC/aMCC, activation of sgACC, pMCC, and PIC are each consistently seen during both negative affect and pain, but not cognitive control (Shackman et al., 2011). A central aspect of justice sensitivity is negative emotional reactions to injustice (Baumert & Schmitt, 2016; Schmitt et al., 1995). Thus, these results are completely consistent with the idea that individuals with high trait self-oriented justice sensitivity would interpret LowReward trials as unfair, and subsequently experience negative emotions as a consequence. This interpretation also may help explain the significant clusters in bilateral hippocampi. A previous study asked participants to solve difficult anagrams, and induced frustration and anger by rudely interrupting participants to ask that they speak louder (Denson, Pedersen, Ronquillo, & Nandy, 2008). Hippocampal activation correlated with a self-report measure of angry rumination. In this interpretation, not only were self-focused individuals upset by the LowReward trials, but this also led them to ruminate on the injustice of the experiment (Baumert & Schmitt, 2009). Moreover, MCC and aINS play important roles in physiological state changes, and their activity here suggests that individuals with higher self-oriented justice sensitivity may have perceived LowReward trials as threatening (Beissner et al., 2013; Benarroch, 1993; Thayer & Lane, 2009).

Finally, several regions showed significant Type x Perspective effects (Table 34), primarily along the frontal midline (Figure 17). Left amygdala, bilateral vmPFC, and more dorsal regions of mPFC all showed a similar pattern, wherein activity was greater for HighReward trials following the Second-Person frame, but greater for LowReward trials following the Third-Person

Frame. This suggests that stimulus-value representations were altered by changing the perspective from which participants evaluated the preceding moral scenarios (Levy & Glimcher, 2012). A cluster in rostral ACC gyrus showed a similar pattern (Figure 18), and suggests that the relative salience of different stimulus types was also altered as a function of the perspective frame (Seeley et al., 2007). This cluster is also interesting because the gyrus of the ACC, just rostral and dorsal to the genu of the corpus callosum, was recently proposed as a functionally distinct region of the ACC (ACCg) which plays a critical role in encoding and maintaining representations of other people's motivations (Apps, Rushworth, & Chang, 2016). In this model, vmPFC and the regions of ACC which line the sulcus (Brodmann areas 24a/b/c) function to integrate cognitive and affective information to update representations in ACCg. Given that this task ostensibly only affected participants it is unclear how these results fit with this interpretation of ACCg functioning, but presents an interesting target for future inquiry.

Finally, both TDT and Dots provided a link between the two questions motivating this dissertation. The OtherFairness x Perspective interaction (Table 26) suggests that after rating the moral permissibility of scenarios that potentially harm a stranger, individuals attach greater significance to monetary distributions which disadvantage a neutral, powerless third-party. During the Dots task, the perspective shift resulted in the reversal of activity in ACC and frontal midline structures for HighReward > LowReward (Table 33). Taken together, these findings suggest that situational cues to injustice directed at the self or others influence neural processing on subsequent tasks.

Overall, Study 5 provided good evidence that other-oriented and self-oriented justice sensitivity are associated with largely distinct neural systems. In economic contexts, the striatum appears to play an important role. Hemodynamic activity in posterior striatum was positively correlated with other-oriented justice sensitivity, but negatively related to self-oriented justice sensitivity (Figure 10). The ventral striatum was revealed in the SelfFair > SelfUnfair main effect contrast (Table 30). Moreover, individuals with higher self-oriented justice sensitivity showed greater striatal response to OtherUnfair compared to OtherFair offers, but only for distributions which were also SelfUnfair (Figure 12). Again, these striatal responses might reflect evaluative coding of the Proposer. In the mutualism framework (Baumard et al., 2013; Baumard & Sheskin, 2015), fairness and morality evolved by partner choice, where individuals choose to associate and engage in interactions with social partners that are cooperative. Thus, identifying those individuals which allocate resources in a fair and equitable manner is central to successful functioning within the group. Thus, striatal activation might be generating prediction error signals regarding the behavior of the Proposer, with distributions that violate the social norm of fairness serving as a cue to be wary of such an individual in the future. This interpretation is further supported by the observed effects in LC and SMA in TDT (Figure 12), as well as MCC and aINS during Dots (Figure 16), which suggest that threat avoidance may emerge as a central facet of self-oriented justice sensitivity.

4.4 Appendix C: Chapter Four Tables

Table 23. Study 5 Demographic Information

	Sample N = 31 Mean (SD) or %
Age	26.6 (8.7)
Percent female	52%
Race/Ethnicity	
White	54%
Black	10%
Hispanic/Latin	6%
Asian	23%
Other or Multiracial	6%

Table 24. Study 5 Dispositional Measures

	Mean (SD)	α
JS-Self	3.7 (1.1)	0.92
JS-Other	4.1 (0.6)	0.88
RS	4.4 (2.6)	0.75
PS	5.6 (4.1)	0.88
IRI-EC	26.4 (4.9)	0.78
IRI-PT	24.2 (4.4)	0.72
IRI-PD	17.4 (5.7)	0.82

Note. JS-Self = Self-oriented justice sensitivity, JS-Other = Other-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity, IRI-EC = Empathic concern, IRI-PT = Perspective taking, IRI-PD = Personal distress. Cronbach's alpha (α) measures internal consistency among individual items, with values above 0.6 considered acceptable.

Table 25. Study 5 Model Parameters for Scenario Ratings

Random Effects	Variance	SD
Participant	0.17	0.41
Type	0.77	0.88
	β	SE
(Intercept)	-0.52†	0.21
Age	-0.02	0.17
Gender (Female)	0.29	0.21
Type (Proscriptive)	0.54†	0.23
Perspective (Third-Person)	0.16	0.23
Type x Perspective	0.03	0.32

Note. Model estimated using linear multilevel model rather than ordinal multilevel model. † $p < .1$ (all p values FDR-corrected).

Table 26. Study 5 Model Parameters for Responses in Distribution Task

Random Effects	Variance	SD
Participant	26.56	5.15
OtherFairness	45.99	6.78
SelfFairness	16.17	5.02
Perspective	3.19	1.79
Fixed Effects	**Log Odds**	**SE**
(Intercept)	8.81***	1.55
Age	0.12	0.83
Gender (Female)	2.11	1.37
OtherFairness (Unfair)	-10.49***	1.65
SelfFairness (Unfair)	-5.07***	1.02
Perspective (Third-Person)	1.88†	0.75
JS-Other	-0.31	1.20
JS-Self	1.03	0.82
OtherFairness x SelfFairness	3.90***	0.63
OtherFairness x Perspective	-0.86	0.70
SelfFairness x Perspective	-0.47	0.68
Perspective x JS-Other	1.17	0.56
OtherFairness x JS-Other	1.07	1.40
Perspective x JS-Self	0.31	0.48
SelfFairness x JS-Self	0.79	0.85
OtherFairness x SelfFairness x Perspective	-0.39	0.79

Table 26, continued

OtherFairness x Perspective x JS-Other	-2.66***	0.54
SelfFairness x Perspective x JS-Self	-0.08	0.34

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. † p < .1, * p < .05, ** p < .01, *** p < .001 (all p values FDR-corrected).

Table 27. Study 5 Model Parameters for Responses in Dots Task

Random Effects	Variance	SD
Participant	4.57	2.14
Type	7.29	2.70
Perspective	0.21	0.46
Fixed Effects	**Log Odds**	**SE**
(Intercept)	3.14***	0.56
Age	-0.51	0.38
Gender (Female)	-0.01	0.50
Type (HighReward)	1.29	0.64
Perspective (Third-Person)	0.34	0.24
JS-Other	-0.38	0.51
JS-Self	-0.26	0.57
RS	-1.02	0.50
PS	-1.37†	0.55
Type x Perspective	0.44	0.61
Type x JS-Other	0.28	0.67
Perspective x JS-Other	-0.10	0.21
Type x JS-Self	-0.11	0.73
Perspective x JS-Self	-0.32	0.21
Type x RS	1.05	0.60
Type x PS	2.00*	0.67
Type x Perspective x JS-Other	0.72	0.51

Table 27, continued

Type x Perspective x JS-Self	1.00	0.54

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all *p* values FDR-corrected).

Table 28. Study 5 Model Parameters for Freerice Responses

Term	Log Odds	SE
(Intercept)	-1.36	0.65
Age	-0.20	0.56
Gender (Female)	1.29	0.85
JS-Other	0.27	0.42
JS-Self	-0.02	0.49

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity

Table 29. Study 5 Whole-Brain Results for OtherUnfair > OtherFair Contrast

		MNI coordinates				
Covariate	Region	x	y	z	Cluster size	T
	R Inferior Temporal	47	-49	-16	11	4.15
	L Inferior Parietal	-31	-60	34	39	4.26
	Precuneus	-6	-56	38	14	4.45
JS-Other						
	L Striatum	-31	-18	-2	12	5.24
JS-Self						
	L Striatum	-24	-18	-2	12	-4.27

Note. R = Right, L = Left, ACC = Anterior cingulate cortex, JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. $P < .001$, k > 10

Table 30. Study 5 Whole-Brain Results for SelfFair > SelfUnfair Contrast

	MNI coordinates				
Region	x	y	z	Cluster size	T
L Inferior Occipital	-31	-91	-9	17	4.03
R Hippocampus	29	-14	-9	16	4.62
Cuneus	-6	-98	6	261	12.02
R Striatum	15	11	-2	27	4.53
Middle Cingulate	5	-28	42	16	4.19

Note. R = Right, L = Left, ACC = Anterior cingulate cortex, JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. $P < .001$, k > 10

Table 31. Study 5 Whole-Brain Results for OtherFairness x SelfFairness Interaction

Covariate	Region	MNI coordinates			Cluster size	T
		x	y	z		
	mPFC	5	56	16	24	-4.20
	R Angular	50	-60	20	35	-4.49
	Precuneus	8	-56	31	82	-5.89
	L Angular	-48	-60	34	36	-5.82
JS-Self						
	Temporal Pole	-31	11	-30	19	-6.49
	Midbrain	5	-35	-16	244	-5.71
	L Striatum	-13	18	-9		-4.82
	L Anterior Insula	-27	18	-5		-4.66
	L Amygdala	-17	0	-16		-4.47
	Cerebellum	-3	-60	-16	35	-4.19
	L Fusiform	-34	-32	-23	24	-4.39
	L Fusiform	-38	-49	-12	31	-4.62
	R Striatum	29	7	-5	52	-4.38
	Precuneus	-3	-49	38	23	-4.02
	SMA	8	-21	60	86	-5.19
	Superior Frontal Sulcus	33	18	45	28	-4.46
	L Precentral	-20	-21	67	15	-4.18

Note. R = Right, L = Left, mPFC = medial prefrontal cortex, SMA = supplementary motor area, JS-Self = Self-oriented justice sensitivity. $P < .001$, $k > 10$

Table 32. Study 5 Whole-Brain Results for OtherFairness x Perspective Interaction

	MNI coordinates				
Region	x	y	z	Cluster size	T
R Postcentral	57	-21	34	56	-5.02
Middle Cingulate	-3	-7	42	16	-3.89

Note. R = Right. *P* < .001, k > 10

Table 33. Study 5 Whole-Brain Results for HighReward > LowReward Contrast

Covariate	Region	MNI coordinates			Cluster size	T
		x	y	z		
	R Lingual	19	-98	-9	17	5.04
	L Striatum	-34	-11	-12	20	3.70
	L Precentral	-41	-11	27	79	4.24
JS-Other						
	L Fusiform	-45	-42	-20	10	3.95
	L Inferior Occipital	-52	-70	-9	18	4.08
	L Midbrain	-13	-32	-2	137	5.04
	L Striatum	-24	0	-9		3.92
	L Insula	-41	0	-9		4.25
	R Striatum	22	7	2	78	4.84
	R Inferior Temporal	54	-53	-5	14	3.50
	R Midbrain	15	-25	-5	23	4.02
	L Striatum	-13	4	2	10	3.82
	R Inferior Frontal	47	32	2	56	4.57
	R Middle Temporal	57	-60	13	40	4.08
	L Inferior Frontal	-45	32	9	95	4.97
	R Calcarine	15	-60	9	34	3.91
	R Supramarginal	64	-28	27	162	5.26
	L Supramarginal	-62	-35	24	168	5.01
	R Middle Occipital	40	-77	31	21	4.30

Table 33, continued

	R Precentral	54	11	38	56	4.38
	R Postcentral	19	-35	56	126	4.33
JS-Self						
	Cerebellum	-3	-53	-20	11	-3.49
	R Hippocampus	22	-11	-12	15	-4.24
	L Hippocampus	-31	-21	-16	10	-3.56
	Rostral Cingulate	1	35	-2	18	-4.66
	R Insula	43	-7	13	30	-4.60
	L Insula	-38	-11	16	26	-4.52
	L Supramarginal Gyrus	-52	-42	27	11	-3.78
	L Precentral	-52	-7	31	95	-4.86
	R Postcentral	57	0	34	45	-4.53
	Middle Cingulate	12	-7	42	28	-3.74
	Precuneus	19	-42	52	19	-3.71
	R Postcentral	29	-28	63	44	-4.20
	L Precentral	-17	-18	70	12	-4.58

Note. R = Right, L = Left, JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. $P < .0025$, k > 10

Table 34. Study 5 Whole-Brain Results for RewardType x Perspective Interaction

	MNI coordinates				
Region	x	y	z	Cluster size	T
Cerebellum	15	-46	-23	33	-4.28
L Amygdala	-20	0	-16	24	-3.86
L vmPFC	-13	46	-5	21	-3.74
Rostral Anterior Cingulate	8	32	2	42	-4.36
R mPFC	15	56	2	15	-4.23
L mPFC	-10	56	6	10	-3.83
L mPFC	-13	46	24	40	-5.09
L Precentral	-20	-25	74	35	-4.08

Note. R = Right, L = Left, mPFC = medial prefrontal cortex, vmPFC = ventromedial PFC. $P < .0025$, $k > 10$

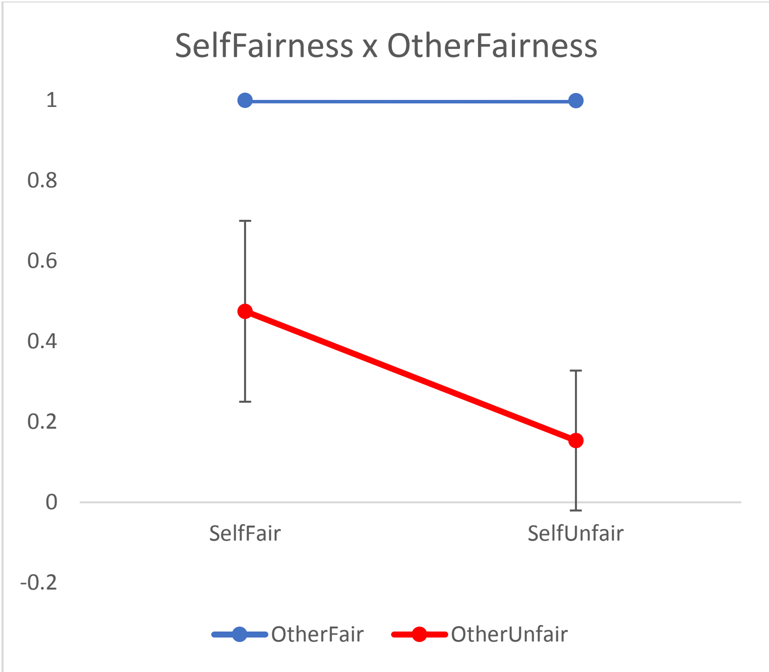


Figure 9. Study 5 SelfFairness by OtherFairness Interaction in the Distribution Task

Mean acceptance rates are shown for visualization purposes. Error bars represent SEM for mean acceptance rates.

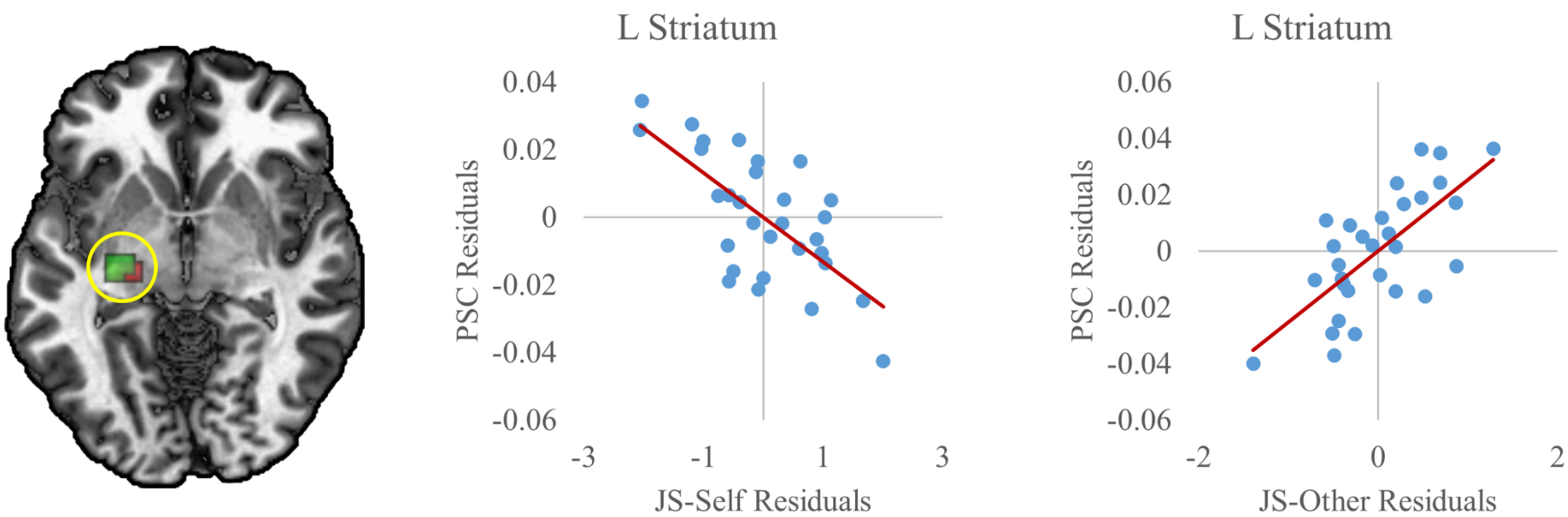


Figure 10. Opposing Influences of Justice Sensitivity for Self and Other in Left Striatum
Whole-brain clusters (left) showing significant ($p < 0.001$, $k = 10$) modulation of self-oriented (JS-Self; red) and other-oriented (JS-Other; green) justice sensitivity. Partial correlations for mean percent signal change, controlling for age, approach/avoidance, and opposite justice sensitivity are shown at right.

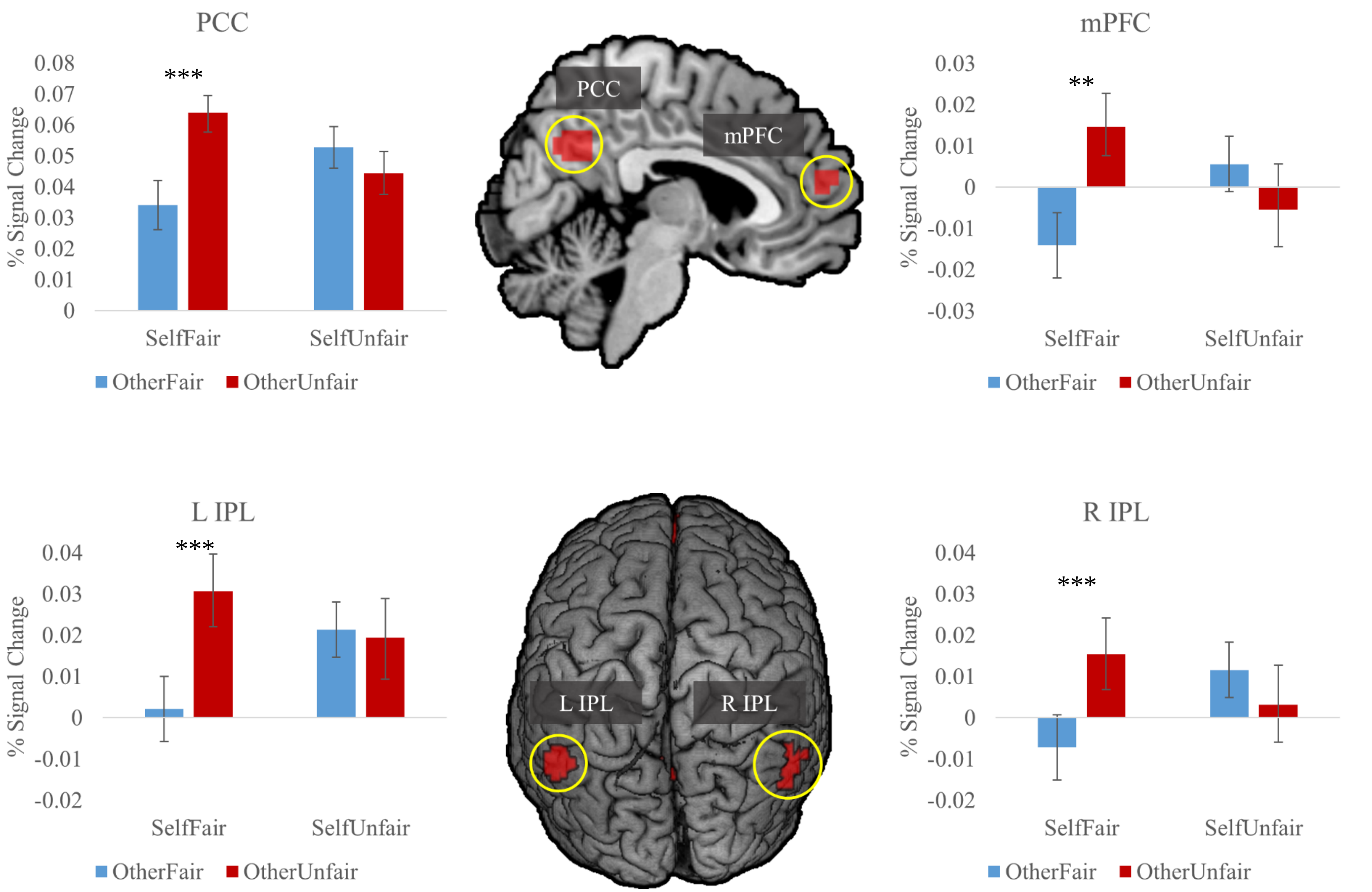


Figure 11. Regions Showing SelfFairness x OtherFairness Interaction
Whole-brain interactions ($p < .001$, $k > 10$) in posterior cingulate cortex (PCC), inferior parietal lobule (IPL) and medial prefrontal cortex (mPFC). ** $p < 0.01$, *** $p < .001$

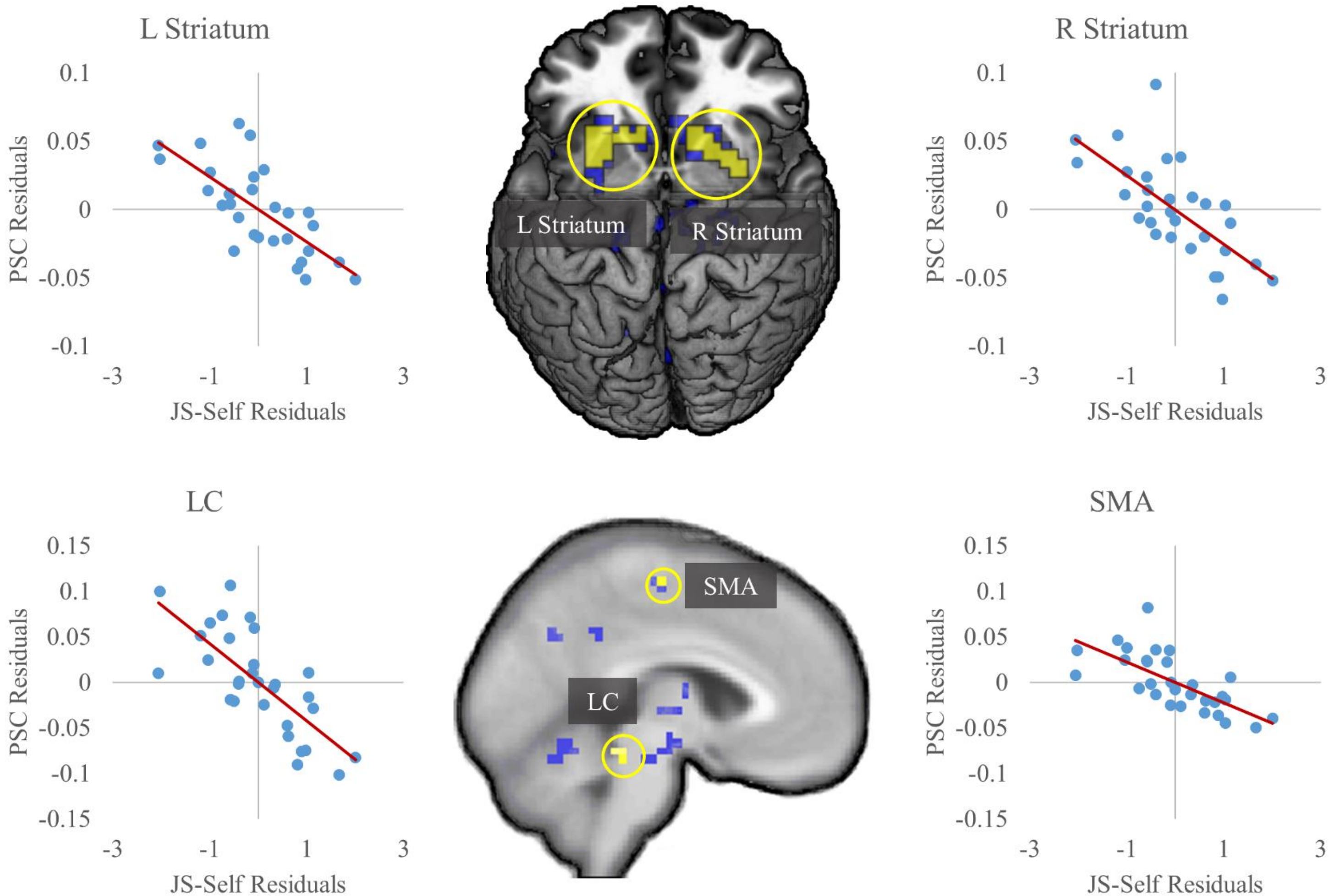


Figure 12. Self-Oriented Justice Sensitivity and Reduced Interaction Effect in Bilateral Striatum and Central Autonomic Network. Significant whole-brain clusters (blue; $p < 0.001$, $k > 10$) for the SelfFairness x OtherFairness interaction showing negative impact of self-oriented justice sensitivity. For visualization, a more conservative threshold was used to obtain a more restricted striatal cluster ($p < 0.0003$). Partial correlations for these restricted clusters, controlling for age, approach/avoidance, and other-oriented justice sensitivity are shown. SMA = Supplementary motor area; LC = Locus coeruleus.

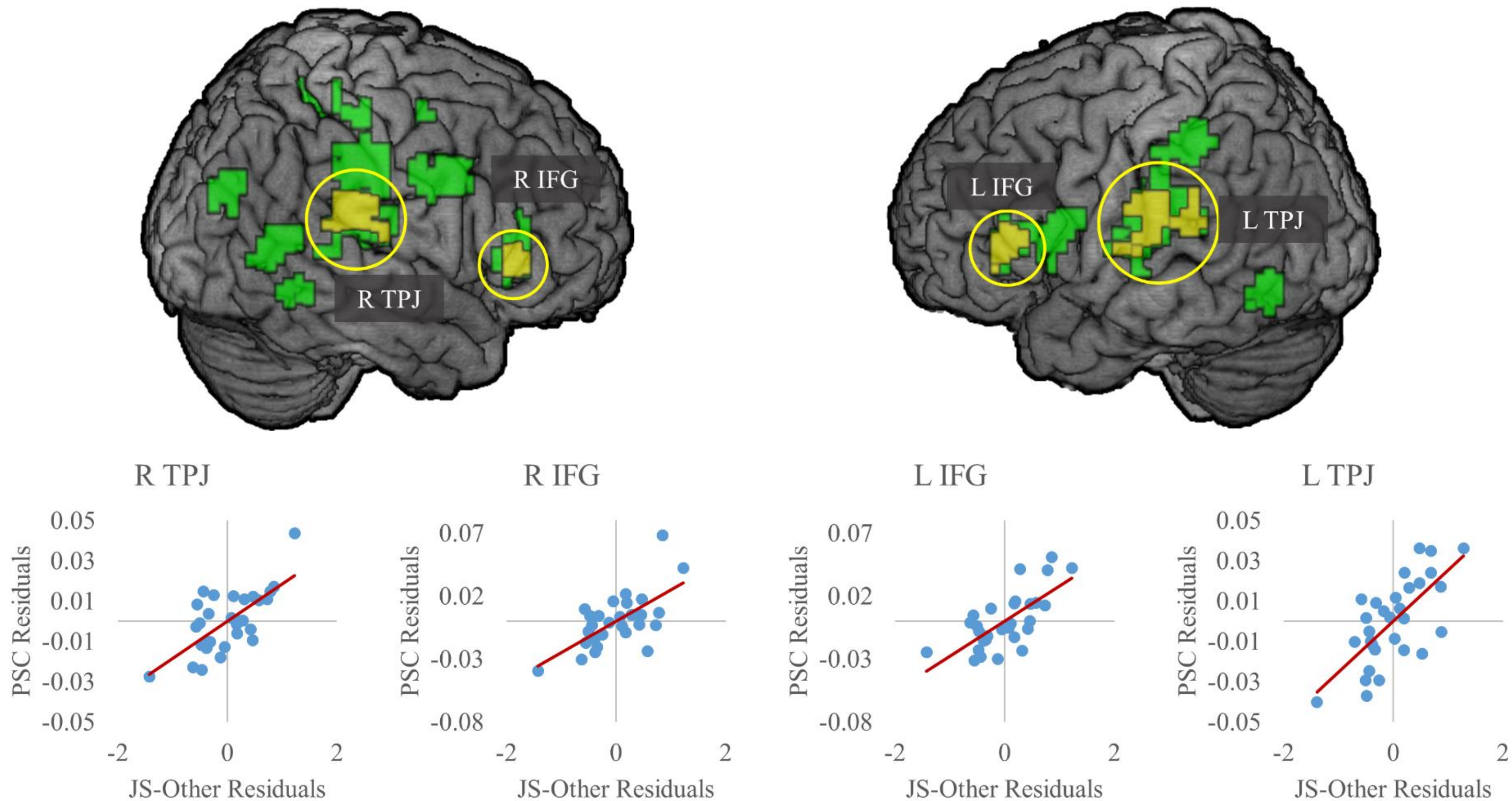


Figure 13. Regions showing positive effect of other-oriented justice sensitivity in Dots task
Whole-brain clusters (HighReward > LowReward) with significant positive relation to other-oriented (JS-Other) justice sensitivity ($p < 0.0025$, $k > 10$), specifically bilateral inferior frontal gyrus (IFG) and temporoparietal junction (TPJ). For visualization, a more conservative threshold was used to obtain restricted clusters (yellow; $p < 0.001$) which were used for partial correlations (bottom).

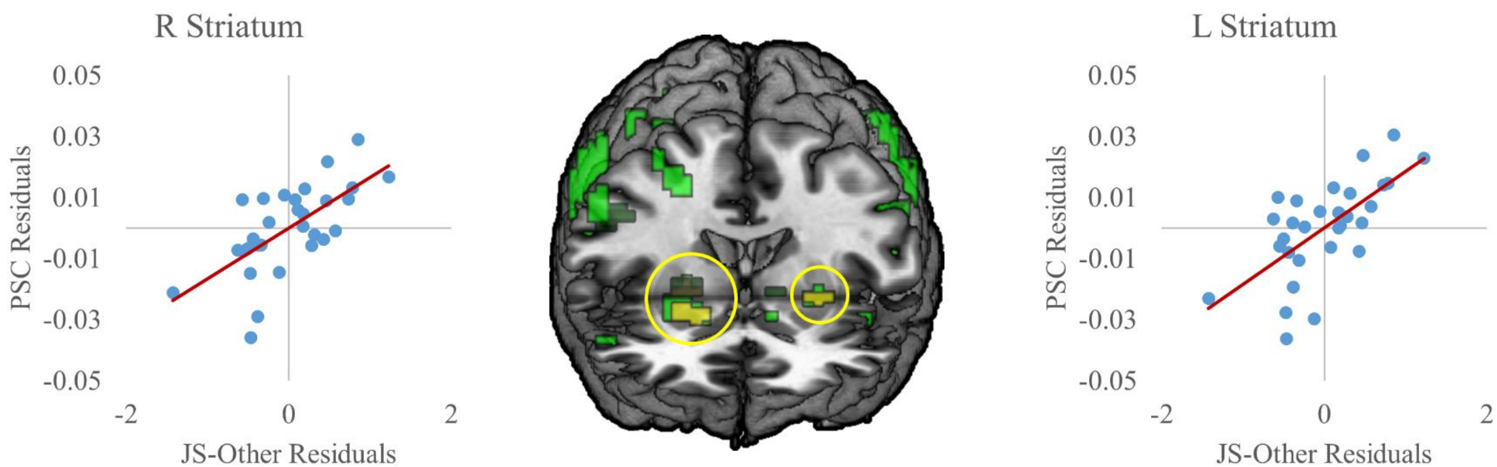


Figure 14. Other-oriented justice sensitivity and increased striatum response for HighReward > LowReward
Whole-brain clusters from HighReward > LowReward showing significant positive relation to other-oriented (JS-Other) justice sensitivity ($p < 0.0025$, $k > 10$). For visualization, a more conservative threshold was used to obtain restricted clusters (yellow; $p < 0.001$) which were used for partial correlations, controlling for age, approach/avoidance motivation, and self-oriented justice sensitivity.

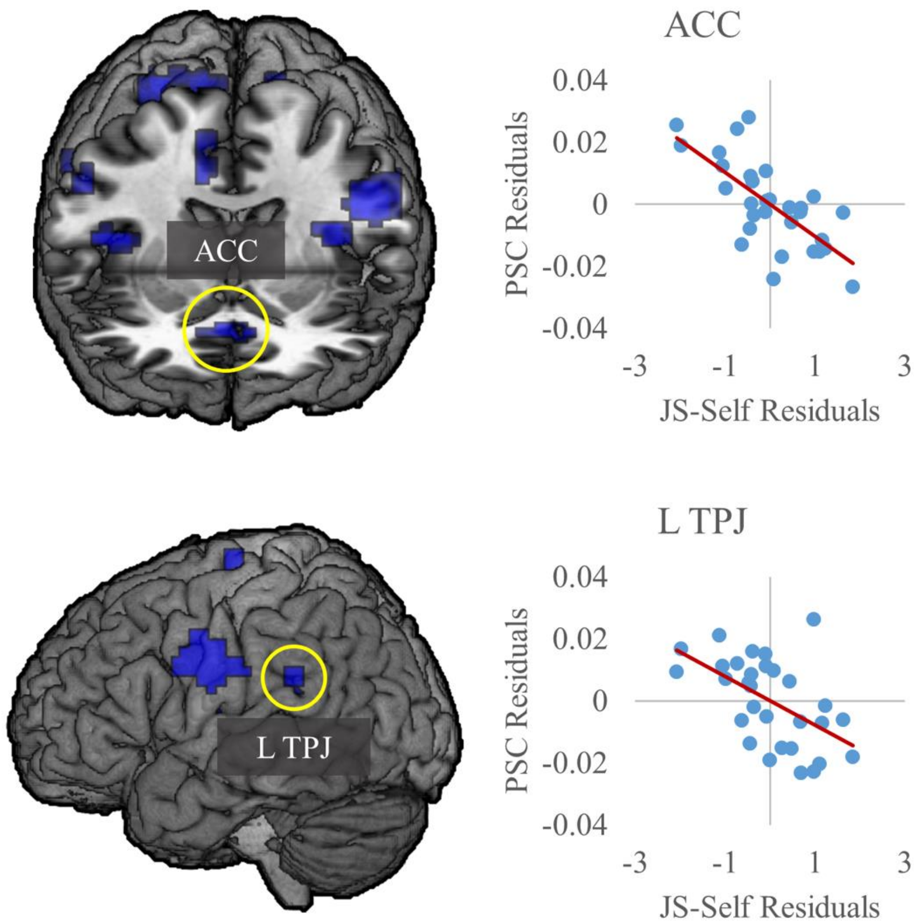


Figure 15. Self-Oriented Justice Sensitivity and reduced signal in ACC and left TPJ. Significant whole-brain clusters from HighReward > LowReward showing significant negative relation with self-oriented (JS-Self) justice sensitivity ($p < 0.0025$, $k > 10$) in anterior cingulate cortex (ACC) and left temporoparietal junction (TPJ). Right: Partial correlations controlling for age, approach/avoidance, and other-oriented justice sensitivity.

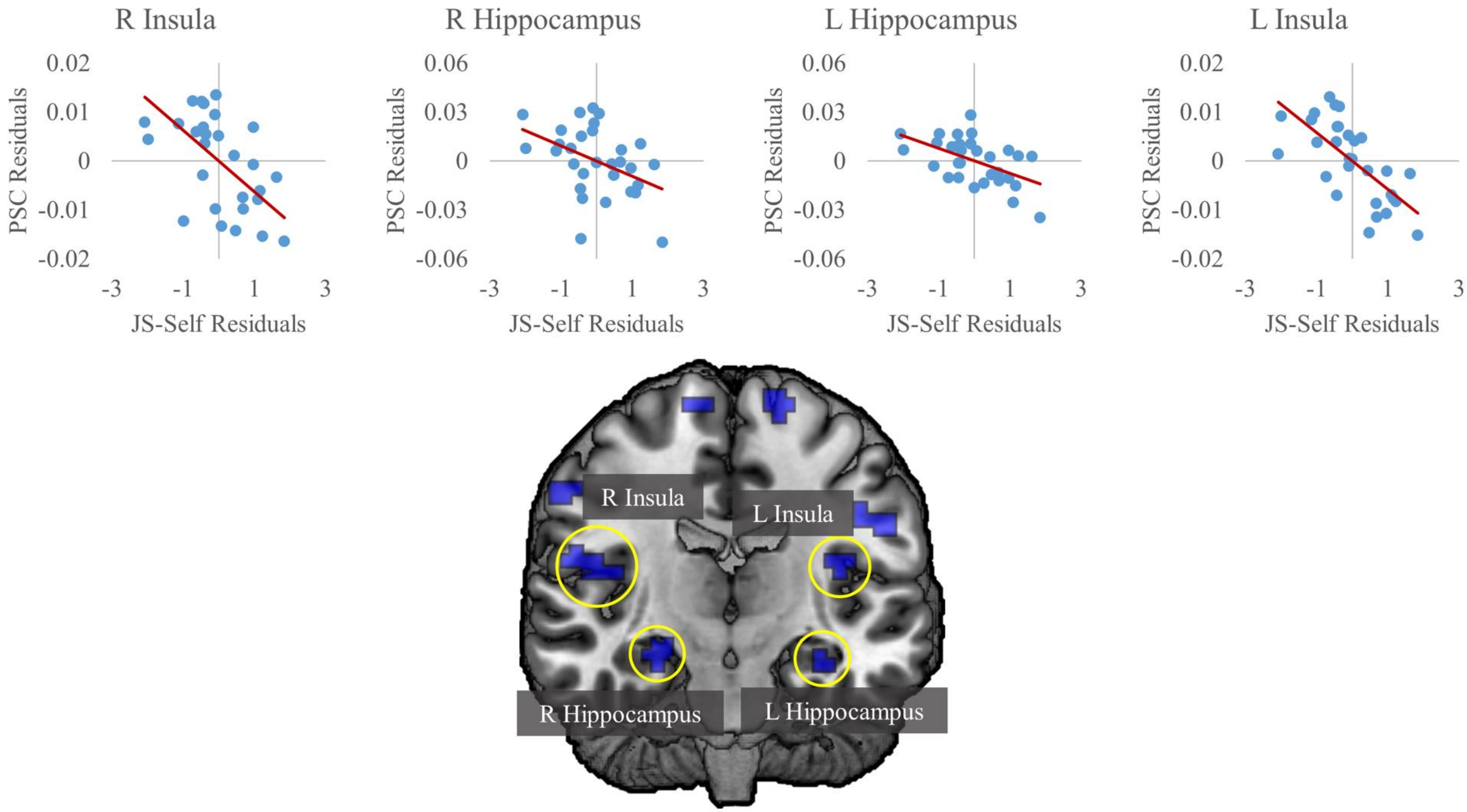


Figure 16. Self-Oriented Justice Sensitivity and Reduced HighReward > LowReward Response in Bilateral Insula and Hippocampi. Significant whole-brain clusters ($p < 0.0025$, $k > 10$) for the HighReward > LowReward showing negative impact of self-oriented justice sensitivity (JS-Self). Top: Partial correlations, controlling for age, approach/avoidance, and other-oriented justice sensitivity.

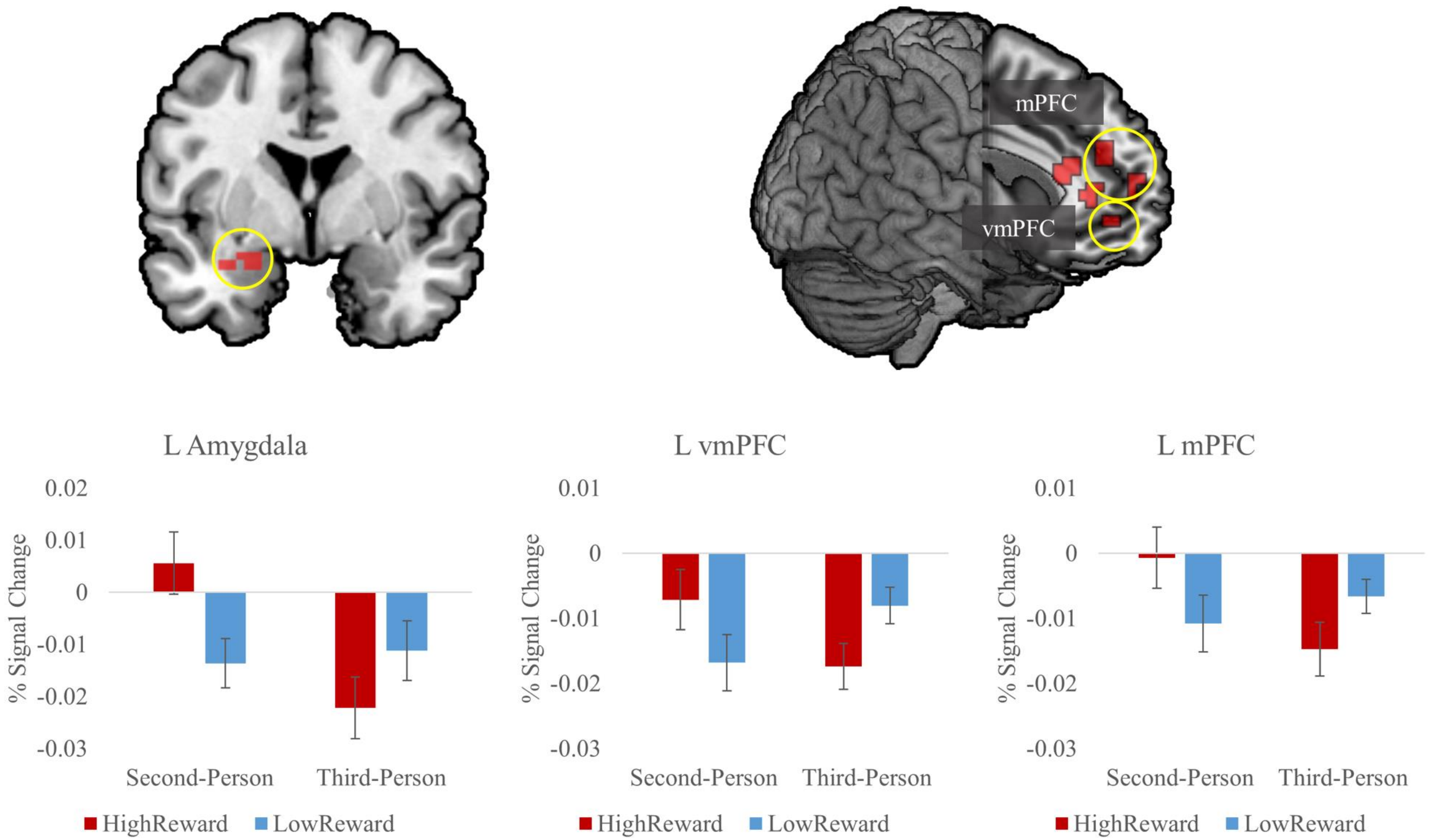


Figure 17. Regions with Perspective x Stimulus Type
Significant whole-brain clusters ($p < 0.0025$, $k > 10$) in amygdala, ventromedial and medial prefrontal cortex (vmPFC, mPFC).

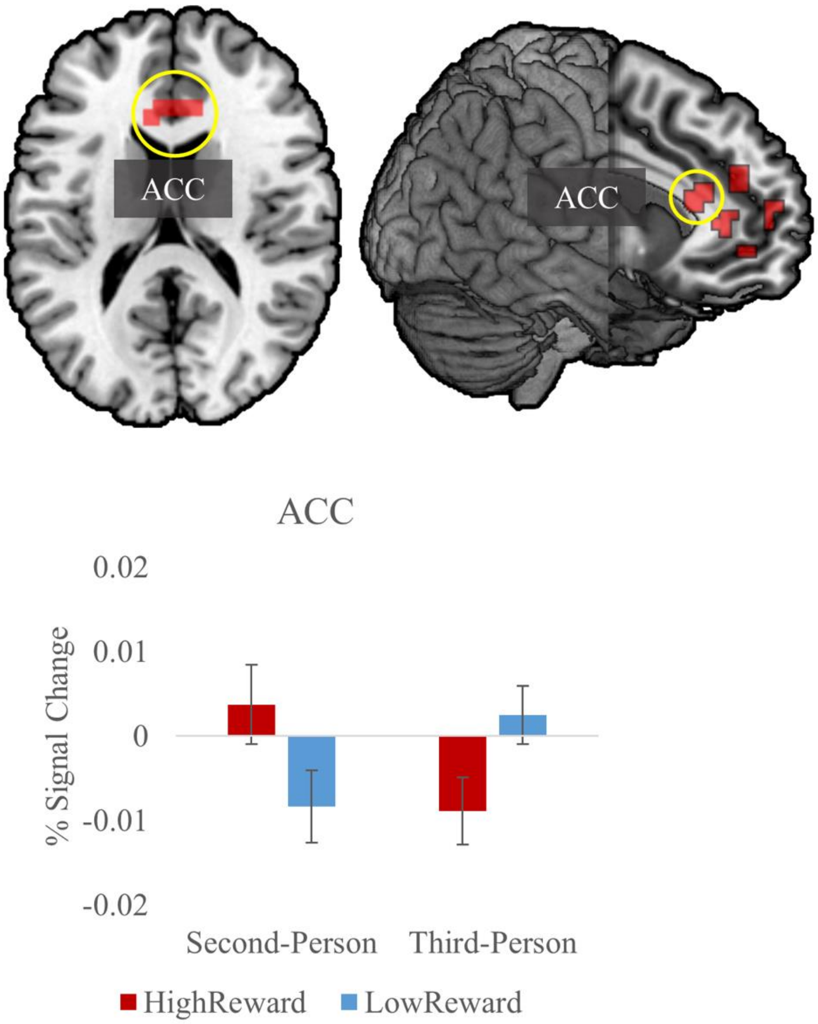


Figure 18. Significant Perspective x Stimulus Type in Anterior Cingulate Cortex
Whole-brain cluster ($p < 0.0025$, $k > 10$) in anterior cingulate cortex (ACC) showing significant Perspective x Stimulus Type interaction. Bottom: Percent signal change from this cluster.

Chapter 5: Effects of Justice Sensitivity on Spatiotemporal Dynamics of Fairness Processing

5.1 Introduction

Study 5 identified associations between justice sensitivity dispositions and activity within several regions important for encoding appetitive and aversive stimuli and facilitating behavioral approach and avoidance, including striatum, amygdala, LC, and aINS, during both the TDT and Dots tasks. This suggests that justice sensitivity may relate to differences in the extent to which justice-related cues influence early sensory encoding of stimuli, with individuals who are highly sensitivity to justice assigning more value to just situations. However, the temporal resolution of fMRI is limited, so Study 5 was ill-equipped to address this questions directly.

Electroencephalography (EEG) involves measuring voltages at the scalp. Event-related potentials (ERPs) are extracted by averaging together multiple traces which are time-locked to the onset of particular stimuli. This allows for the investigation of online neural processing with millisecond temporal resolution and makes the ERP technique well-suited to test whether individual differences in justice sensitivity influence early sensory encoding or later deliberative stages. Moreover, decades of research in cognitive control and affective processing have identified several ERP components which are reliably associated with different aspects of information processing (Folstein & Petten, 2008; Hajcak et al., 2010; Luck, 2014; Luck & Kappenman, 2012).

For instance, investigations into affective processing distinguish between early (100 – 200 ms), middle (200 – 300 ms), and late (> 300 ms) ERPs (Olofsson, Nordin, Sequeira, & Polich, 2008). Early and middle latency components are argued to reflect relatively automatic and obligatory attention capture (Macnamara, Foti, & Hajcak, 2009; Weinberg & Hajcak, 2010), and show greater amplitudes to emotional images, compared to neutral images. For instance, negative deflections over frontocentral sites manifest when observing others in physical pain, irrespective of task demands (Cheng, Hung, & Decety, 2012; Decety, Yang, & Cheng, 2010; Fan & Han, 2008; Perry, Bentin, Bartal, Lamm, & Decety, 2010). Many investigations into emotion processing have be influenced by the negativity bias framework, which argues that a central role of emotion is to rapidly identify aversive and dangerous stimuli in the environment (Cacioppo, Gardner, & Berntson, 1999; Crawford & Cacioppo, 2002; Ohman, Flykt, & Esteves, 2001). This led to an initial suspicion that early and middle latency components should show enhanced amplitudes for aversive compared to appetitive stimuli. However, a number of studies have instead found greater amplitudes for positively valenced stimuli in the earlier posterior negativity (EPN), compared to negatively valenced stimuli, whether those stimuli are affective pictures or morally laden scenarios, both in adults and young children (Cowell & Decety, 2015b; Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Keil et al., 2002; Schupp, Markus, Weike, & Hamm, 2003; Weinberg & Hajcak, 2010; Yoder & Decety, 2014a).

Between 200 and 350 ms after the presentation of a stimulus, a negative scalp deflection often emerges over frontocentral sites. Depending on the context and research field, this component is referred to as the N2, error-related negativity (ERN), feedback-related negativity (FRN), or medial frontal negativity (MFN). It has been implicated in novelty, expectancy violation, and cognitive control (Botvinick, Cohen, & Carter, 2004; Cohen, Botvinick, & Carter, 2000; Folstein & Petten, 2008), and source estimation techniques place the neural generator for this scalp frontocentral negativity within the ACC (Hajcak, Moser, et al., 2005). As discussed in *Chapter 1* and *Chapter 4*, the ACC plays a critical role in a broad range of tasks, potentially by integrating information from other regions in order to effectively allocate attentional and cognitive resources (Harsay et al., 2012; Shackman et al., 2011; Shenhav, Cohen, & Botvinick, 2016). In other words, there is good evidence to suggest that frontal N2 amplitudes often reflect the perceived motivational significance of a stimulus (Hajcak, Moser, et al., 2005). This has not escaped the notice of neuroeconomists, and several studies have demonstrated enhanced N2 amplitudes for unfair compared to fair offers in the Ultimatum Game (Boksem & De Cremer, 2010; Massi & Luhmann, 2015; Mussel, Hewig, Allen, Coles, & Miltner, 2014; Wang, Li, Li, Wei, & Li, 2016; Zanolie, De Cremer, Güroğlu, & Crone, 2015).

A few previous studies have investigated ERP responses to fair and unfair offers in a three-party ultimatum games, and each focused on middle latency ERP components. The first found a main effect of SelFairness, with enhanced frontal negativity (FCz: 280 – 360 ms) for offers which were unfair to the self (Alexopoulos et al., 2012). There was not a significant main effect for OtherFairness or an interaction, and the authors interpreted this to mean that early ERP indices are responsive to unfairness, but only with respect to the self. However, another study found no main effects on frontal negativities (Cz: 220 – 320 ms), but instead a significant interaction, with response to SelfUnfair – OtherUnfair greater than the other three types of distributions (Alexopoulos et al., 2013). Moreover, there was a significant correlation between the perpetrator subscale of other-oriented justice sensitivity and MFN amplitudes in the SelfFair - OtherUnfair minus SelfFair – OtherFair difference wave. Importantly, these ERP effects were not present when participants knew that offers were made by a computer, again pointing towards the social nature of these fairness responses. Finally, one recent study with male university students (aged 18-27 years) found that when participants believed that the affected observer was an attractive female, they were more willing to accept unfair offers, and the frontal negativity (six frontocentral sites: 290 – 370 ms) was no longer significantly different between any combination of SelfFairness and OtherFairness (Ma & Hu, 2015). However, when the observer was less attractive, there was a pronounced MFN for SelfUnfair – OtherFair offers. Thus, middle latency frontal ERPs represent a good candidate for influences of both perspective framing and justice sensitivity dispositions.

Later ERPs are thought to be influenced by cognitive appraisal and other deliberative processes, but also subjective arousal (Macnamara et al., 2009; Olofsson et al., 2008; Smith, Weinberg, Moran, & Hajcak, 2013; Weinberg & Hajcak, 2010). The P3 is a very well-studied component which is positive over most of the superior surface of the scalp. It is separable into the P3a and P3b, with the former reflecting attentional processing in frontal cortex, and the latter reflecting memory processing in the parietal and temporal cortices (Polich, 2007). P3 is greater for rare stimuli in oddball tasks, but also shows greater amplitudes for stimuli with emotional content. Since this effect appears even when emotional content is task-irrelevant, some have suggested that emotional stimuli are "natural targets" (Macnamara et al., 2009), which fits with the proposed role of emotion in guiding attention towards behaviorally relevant stimuli (Lang & Davis, 2006). Following the P3 complex is the late positive potential (LPP), thought to capture the influence of attentional process after initial sensory encoding (Cacioppo, Crites, & Gardner, 1996). Rather than a simple positive-negative distinction, LPP amplitudes are sensitive to fine-grained characteristics of stimuli (Macnamara et al., 2009) as well as social context (Bamford et al., 2015; Y. Wu, Leliveld, & Zhou, 2011) and whether participants find content of a stimulus to be arousing (Cuthbert et al., 2000).

Thus, the ERP technique provides an opportunity to directly test several of the proposed mechanisms of justice sensitivity (Baumert & Schmitt, 2016). First, the activation threshold for injustice concepts is lower for individuals with higher dispositional justice sensitivity, meaning that these individuals require fewer cues of injustice to perceive a situation as unjust (Baumert &

Schmitt, 2009; Schmitt et al., 1995). Second, justice sensitivity includes affective reactivity, with perceived injustice evoking stronger negative emotional responses (Mohiyeddini & Schmitt, 1997; Schmitt & Mohiyeddini, 1996). To date, studies which have supported these mechanisms have relied on behavioral responses and self-report metrics (Baumert, Schlösser, et al., 2014; Fetchenhauer & Huang, 2004; Gollwitzer & Rothmund, 2011; Lotz et al., 2011). Thus, Study 6 investigated these claims using well-established ERP indices of attention capture (e.g. EPN/N1 and N2), as well as subjective arousal (e.g. LPP).

Individuals with higher other-oriented justice sensitivity were expected to show enhanced N2 and LPP amplitudes for offers which were unfair for the observer. Similarly, self-oriented justice sensitivity was expected to predict greater N2 and LPP amplitudes for SelfUnfair distributions, especially when distributions were SelfUnfair-OtherFair. At the same time, self-oriented justice sensitivity was expected to predicted greater EPN amplitudes for SelfFair offers, because they are expected to find these distributions more pleasant. Moreover, perspective framing is expected to bias individuals to attend to more towards the fairness of the corresponding target, and so perspective was predicted to also enhance N2 and LPP amplitudes for unfair offers directed at the participant or observer corresponding to the preceding frame.

EEG measures of dishonesty and deceptive behavior have primarily been of two types. Some build on the Guilty Knowledge Test (GKT), and embed a target item (e.g. participants' own birthday) in sequences of unrelated "dummy" stimuli and ask participants to feign ignorance of the target item. Studies employing the GKT leverage the fact that the P3 is greater for targets and novel stimuli and thus shows a greater response to the rare target of deception. In these studies, both peak and mean amplitude of the P3 is reliably enhanced these target item (A. R. Miller & Rosenfeld, 2004). Others studies have used a similar approach to those described in

Chapter 4, where individuals respond dishonestly to whether or not a practiced stimulus set is remembered or not. Here, the role of the ACC in conflict monitoring is again implicated. Frontocentral N2 amplitudes are greater for lies, (Hu, Wu, & Fu, 2011; H. Wu, Hu, & Fu, 2009). Moreover, later frontal negativities (400 – 500 ms) have been observed both for lies and truth told with the intent to deceive (Carrión, Keenan, & Sebanz, 2010). Thus, decisions to respond dishonestly during the Dots task were expected to be associated with enhanced frontal N2 amplitudes, as well as later frontal differences. Self-oriented justice sensitivity was expected to predict enhanced P3 amplitudes for HighReward compared to LowReward trials because they were expected to be more motivated to seek out the rarer HighReward trials. Conversely, because other-oriented individuals were expected to be less tempted to respond dishonestly to LowReward trials, other-oriented justice sensitivity was expected to predict reduced frontal N2 amplitudes for the LowReward – HighReward difference wave.

5.2 Methods

5.2.1 Participants

In total, 40 adults participated in the EEG study. Four participants failed to achieve 90% accuracy on the Dots practice, and were excluded from all analyses (see Table 35 for demographic information). Participants were compensated with $20, plus whatever they earned during the Dots Task (*M* = $5.70) with a combination of cash and gift cards. Exclusion criteria included reported use of psychoactive medication or prior head trauma. Study materials and procedures were approved by the Institutional Review Board at the University of Chicago. All participants provided informed written consent.

5.2.2 Dispositional Measures

After phone screening, and approximately one week prior to scanning, participants were emailed an individualized link to Qualtrics to complete dispositional measures. Self-oriented and other-oriented justice sensitivity were assessed using the full-length JSI. The SPSRQ-SF and IRI were used to measure dispositional reward and punishment sensitivity and empathic dispositions, respectively. Descriptive statistics and reliability information for these measures are shown in Table 36.

5.2.3 Stimuli and Tasks

Moral Framing Scenarios. Following the same general protocols as Study 5, participants rated the moral permissibility of half of the scenarios in either the Second-Person or Third-Person perspective (order counter-balanced between participants). This was followed by one run each of TDT and Dots. After a short break, participants rated the second half of the scenarios from the other perspective. Participants provided responses by using the arrows keys on a keyboard to move a bar along a 7-point permissibility scale (1 = Completely, 4 = Somewhat, 7 = Not at all).

Distributional Fairness. The TDT was adapted from the MRI version in three ways. First, stimuli were modified slightly to ensure image fidelity and compensate for the increased native resolution of the display monitor compared to the MRI back-projection system. Second, the MRI version of the TDT used 16 trials for each SelfFairness x OtherFairness x Perspective combination. While sufficient for assessing differences in hemodynamic response, this trial count was too low for EEG (Moran, Jendrusina, & Moser, 2013). Thus, TDT trial counts were doubled (32 trials per condition, total 128 trials/run). Finally, participants gave responses in

Study 6 using both hands on the keyboard ("a" and "l" keys) as in Study 1. The "Accept" and "Reject" mappings were counter-balanced between participants.

Antisocial Behavior. The EEG version of the Dots Task was identical to that used in Study 5, except that responses were given using both hands ("a" and "l" keys) as with the EEG version of the TDT. Each run contained 100 stimuli (60 LowReward and 40 HighReward). The stimulus-reward mapping and stimulus set used for runs one and two were counter-balanced between participants. Participants were not told about the payment schedule until after completion of the practice trials, and participants never saw an individual Dots display twice.

Prosocial Behavior. Study 6 followed the same procedure as Study 5, with a link explaining Freerice being emailed to participants after all participants had completed the study.

5.2.4 EEG Acquisition and Processing

Stimuli were presented via a 23" Samsung S23A700D monitor with 2 ms response rate with a native display resolution of 1920 x 1080 at 60 Hz refresh rate. Participants viewed stimuli while seated in a chair approximately 80 cm from the monitor. EEG data were collected using a 32-channel BrainVision ActiChamp measurement system. Electrode impedances were kept within the manufacturer's recommendations (max 30 kΩ). Scalp voltages were digitized at 2,000 Hz and referenced to Cz. EEG data were processed offline using BrainVision Analyzer (BVA; Version 2.1). First, Cz was recovered, and data were re-referenced to the average of all channels. Next, low-pass (0.1 Hz, 12 dB/octave), high-pass (30 Hz, 24 dB/octave), and notch (60 Hz) filtering was performed using a Butterworth zero phase IIR filter. Data were then downsampled to 256 Hz before using independent component analysis (ICA) to remove ocular artifacts. As implemented in BVA, restricted infomax ICA was performed, and ostensible blink components were automatically identified by time-course activity and scalp distribution. After visual inspection, blink components were removed and the remaining components were back-projected into sensor-space. Following ocular correction, semi-automatic procedures were used to identity periods of artifact within individual channels. In addition to standard procedures (max gradient = 100 μV/ms; max difference within 200 ms = 200 μV; low activity cutoff = 0.5 μV) data were visually inspected for artifacts. Any channel with more than 20% of its data marked as artifact was removed and not included in subsequent analyses.

After data cleaning, segments were extracted beginning 200 ms before stimulus onset and continuing until 1000 ms after stimulus onset. Segments were time-locked to the offer slide in the TDT and the dots display during the Dots task. After segmentation, baseline correction was applied using the average of the 200 ms preceding stimulus onset. Finally baseline-corrected trials were averaged together to create individual ERPs.

5.2.5 Data Analysis

5.2.5.1 Analysis of Behavioral Data

The analysis of behavioral responses in Study 6 was identical that used in Study 5. Briefly, scenario responses were modeled using linear multilevel modeling, averaging within each level of Type x Perspective an assuming an interval scale. TDT decisions (0 = Reject, 1 = Accept) were modeled using a generalized multilevel models with a logit link and random slope terms for each level of OtherFairness, SelfFairness, and Perspective. Dots responses (0 = Incorrect, 1 = Correct) were modeled with the same approach, including random slope terms Type and Perspective. A logistic regression was used to model respond to the Freerice follow-up email.

5.2.5.2 ERP analysis

Based on previous literature and visual inspection of time-course voltages and scalp maps, frontal and posterior midline electrodes were chosen for analysis (Fz and Oz, respectively). Mean amplitudes within each time window (N1: 100 - 175 ms; N2: 150 - 250 ms; P3: 300 - 500 ms; LPP: 600 - 1000 ms) were extracted for each stimulus type in each perspective. Only trials to which participants provided a response were included in the analysis. All participants had at least 16 artifact-free segments for each condition in TDT and Dots. Prior to statistical analysis, averages within each time window in each condition were examined, and any individual whose mean amplitude was more than three standard deviations above or below the mean for that condition was excluded from analysis. This threshold excluded 9 participants from the TDT ERP analysis and 5 participants from the Dots ERP analysis.

The specific ERPs elicited by TDT trials (N1 at Oz, N2/EPN at Oz, N2/MFN at FZ, P3 at Oz, and LPP at Oz) were assessed using a 2 (SelfFairness) x 2 (OtherFairness) x 2 (Perspective) repeated measures ANOVAs. In order to assess the potential impact of self-oriented and other-oriented justice sensitivity on ERP response, I first computed difference waves for the main effect of SelfFairness (SelfUnfair – SelfFair) and OtherFairness (OtherUnfair – OtherFair), the SelfFairness x OtherFairness interaction ($\text{SelfUnfair}_{\text{OtherUnfair-OtherFair}}$ – $\text{SelfFair}_{\text{OtherUnfair-OtherFair}}$), and the SelfFairness x Perspective and OtherFairness x Perspective interactions (Second-Person – Third-Person). Each difference ERP was then regressed on justice dispositions. Age, gender, and reward and punishment sensitivity were included as covariates of no interest.

Similarly, Dots ERPs (N1 at Oz, N2/EPN at Oz, N2/MFN at Fz, P3 at Oz, and LPP at Fz), were interrogated using 2 (Stimulus Type) x 2 (Perspective) repeated measures ANOVAs. The influence of justice dispositions was assessed in the same fashion as for the TDT task, using the HighReward – LowReward and $\text{SecondPerson}_{\text{HighReward-LowReward}}$ – $\text{ThirdPerson}_{\text{HighReward-LowReward}}$ difference waves. Again, age, gender, and reward and punishment sensitivity scores were modeled as covariates of no interest.

5.3 Results and Discussion

5.3.1 Behavioral Responses

Scenario Ratings. The linear multi-level model including demographic and task variables was a significant fit for the data ($X^2(5) = 28.2$, $p < .001$; Table 37). Justice dispositions did not significantly improve upon that model ($p > .8$). Participants rated proscriptive violations as worse the prescriptive violations ($\beta = 0.72$, $p = .007$).

Distributional Fairness. Table 38 show the model parameters for responses during the TDT. Younger participants were marginally less likely to accept any offer (OR = 0.04, $p = .097$). As in Study 5, there were significant main effects of OtherFairness (OR = 0.00, $p < .001$) and SelfFairness (OR = 0.00, $p < .001$), as well as a significant OtherFairness x SelfFairness interaction (OR = 18046.76, $p < .001$; Figure 19). The mutual fairness distribution was more likely to be accepted than any of the other three distributions (*p*s < .001). SelfFair-OtherUnfair offers were more likely to be accepted than SelfUnfair-OtherUnfair offers ($p = .006$), and SelfUnfair-OtherFair offers were marginally more likely to be accepted than mutually unfair offers ($p = .093$). SelfFair-OtherUnfair and SelfUnfair-OtherFair rejection rates were not significantly different ($p > 0.3$).

Ethical Behavior. Accuracy during the Dots task was marginally more likely following the Third-Person Frame (OR = 10.38, $p = .067$). No other predictor variable was individually significant after correction for multiple comparisons (Table 39).

Prosocial Behavior. Freerice response was not significantly predicted by any of the terms in the model (Table 40).

5.3.2 Electrophysiological Responses

Distributional Fairness. The posterior N1 was not significantly modulated by task variables after correction for multiple comparison (all *ps* > 0.2). There was a significant main effect of SelfFairness on the EPN ($F(1,25) = 41.74$, $p < .001$, $\eta_p^2 = .030$), posterior P3 ($F(1,25) = 24.37$, $p < .001$, $\eta_p^2 = .029$), and posterior LPP ($F(1,25) = 28.45$, $p < .001$, $\eta_p^2 = .040$). Amplitudes for SelfFair stimuli were more negative than SelfUnfair stimuli during each time period (Figure 20; *ps* < .001). None of the terms in the ANOVA reached significance for frontal scalp potentials (all *p*s > .2). The enhanced EPN for SelfFair related to SelfUnfair distributions matches previous investigations which have found greater EPN deflections for pleasant images compared to unpleasant images (Cuthbert et al., 2000; Keil et al., 2002; Schupp et al., 2003; Weinberg & Hajcak, 2010). It also mirrors previous studies showing enhanced EPN amplitudes for dyadic interactions depicting helpful interaction compared to harmful interactions (Cowell & Decety, 2015b; Yoder & Decety, 2014a). Thus, greater EPN deflections for SelfFair offers is consistent with individuals perceiving situations where they receive more money as more pleasant.

The lack of a significant frontal N2 (MFN) was surprising, given that this component is the primary target of most electrophysiological investigations of ultimatum-style games and has previously been implicated in three-party ultimatum tasks (Alexopoulos et al., 2013, 2012). In these previous studies, experimenters stressed that distributions were being made by another real person, an aspect that was absent here and may have contributed to participants perceiving this task in a different fashion. It is worth noting that the two previous studies using three-party ultimatum distributions do not appear to have corrected for multiple comparisons. If that approach had been adopted here, there would still be no significant main effect of fairness on MFN amplitude (all uncorrected $p > .13$), but the analysis would have identified a significant positive correlation between self-oriented justice sensitivity and MFN for the SelfFair – SelfUnfair difference wave (uncorrected $p = 0.037$). However, after correction for multiple comparisons, neither self-oriented nor other-oriented justice sensitivity dispositions significantly predicted activity for any of the planned contrasts in any time window.

Ethical Behavior. Though several ERPs demonstrated sensitivity to stimulus type and the Type x Perspective interaction, none of these effects survived correction for multiple comparisons (all *ps* > 0.2). As in the distribution task, dispositional self-oriented and other-oriented justice sensitivity scores did not significantly predict any of the ERP difference waves.

Taken together, results from Study 6 suggest that during three-party distribution games, neural processing in both middle and late stages is impacted more by whether distributions are fair for the self than by whether they are fair for a neutral observer. Moreover, Study 6 found no support for a link between dispositional justice sensitivity and ERP indices of attention capture or subjective arousal. There was also no evidence to suggest that perspective framing influenced the spatiotemporal neural dynamics of fairness processing or ethical behavior. Future studies with more participants or trials per condition will be needed to clarify whether justice dispositions or perspective-framing might produce small effects which were missed here.

5.4 Appendix D: Chapter Five Tables

Table 35. Study 6 Demographic Information

	Sample N = 36 Mean (SD) or %
Age	23.2 (5.6)
Percent female	64%
Race/Ethnicity	
White	44%
Black	14%
Hispanic/Latin	11%
Asian	19%
Other or Multiracial	11%

Table 36. Study 6 Dispositional Measures

	Mean (SD)	α
JS-Self	3.7 (1.1)	0.84
JS-Other	4.2 (0.9)	0.94
RS	4.5 (2.3)	0.74
PS	6.7 (4.2)	0.79
IRI-EC	27.0 (5.2)	0.88
IRI-PT	24.4 (3.7)	0.61
IRI-PD	18.0 (5.7)	0.69

Note. JS-Self = Self-oriented justice sensitivity, JS-Other = Other-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity, IRI-EC = Empathic concern, IRI-PT = Perspective taking, IRI-PD = Personal distress. Cronbach's alpha (α) measures internal consistency among individual items, with values above 0.6 considered acceptable.

Table 37. Study 6 Parameters for Linear Model of Scenario Responses

Random Effects	Variance	SD
Participant	0.10	0.32
Type	0.70	0.84
Fixed Effects	**β**	**SE**
(Intercept)	-0.82***	0.19
Age	0.12	0.09
Gender (Female)	0.51	0.19
Type (Proscriptive)	0.72**	0.20
Perspective (Third-Person)	0.40	0.20
Type x Perspective	-0.26	0.28

Note. Model estimated using linear multilevel model rather than ordinal multilevel model. ** p < .01, *** p < .001 (all p values FDR-corrected).

Table 38. Study 6 Model Parameters of Responses in Distribution Task

Random Effects	Variance	SD
Participant	52.91	7.27
OtherFairness	25.18	5.02
SelfFairness	57.18	7.56
Perspective	4.20	2.05
Fixed Effects	**Log Odds**	**SE**
(Intercept)	16.25***	2.04
Age	-3.26†	1.22
Gender (Female)	-1.15	2.59
OtherFairness (Unfair)	-11.93***	1.18
SelfFairness (Unfair)	-14.56***	1.53
Perspective (Third-Person)	-0.40	0.77
JS-Other	1.29	1.34
JS-Self	-0.02	1.31
OtherFairness x SelfFairness	9.80***	0.66
OtherFairness x Perspective	0.05	0.69
SelfFairness x Perspective	0.25	0.66
Perspective x JS-Other	-0.72	0.48
OtherFairness x JS-Other	-1.61	0.96
Perspective x JS-Self	0.72	0.57
SelfFairness x JS-Self	-0.01	1.34
OtherFairness x SelfFairness x Perspective	-0.87	0.80

Table 38, continued

OtherFairness x Perspective x JS-Other	-0.11	0.37
SelfFairness x Perspective x JS-Self	-0.76	0.43

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$ (all p values FDR-corrected).

Table 39. Study 6 Model Parameters for Responses during Dots Task

Random Effects	Variance	SD
Participant	14.50	3.67
Type	1.56	1.25
Perspective	17.71	4.21
Fixed Effects	**Log Odds**	**SE**
(Intercept)	1.98*	0.72
Age	0.19	0.32
Gender (Female)	0.06	0.37
Type (HighReward)	0.42	0.32
Perspective (Third-Person)	2.34*	0.82
JS-Other	1.39	0.74
JS-Self	0.50	0.70
RS	0.09	0.24
PS	0.21	0.20
Type x Perspective	-0.10	0.43
Type x JS-Other	0.54	0.46
Perspective x JS-Other	-1.05	0.84
Type x JS-Self	-0.37	0.35
Type x RS	-0.27	0.81
Type x PS	0.58	0.39
Perspective x JS-Self	0.25	0.30
Type x Perspective x JS-Other	-0.42	0.42

Table 39, continued

Type x Perspective x JS-Self	-0.18	0.37

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity, RS = Reward sensitivity, PS = Punishment sensitivity. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 40. Study 6 Model Parameters for Freerice Response

Term	Log Odds	SE
(Intercept)	-1.95	0.83
Age	-0.21	0.39
Gender (Female)	2.11	0.96
JS-Other	-0.28	0.44
JS-Self	0.72	0.48

Note. JS-Other = Other-oriented justice sensitivity, JS-Self = Self-oriented justice sensitivity

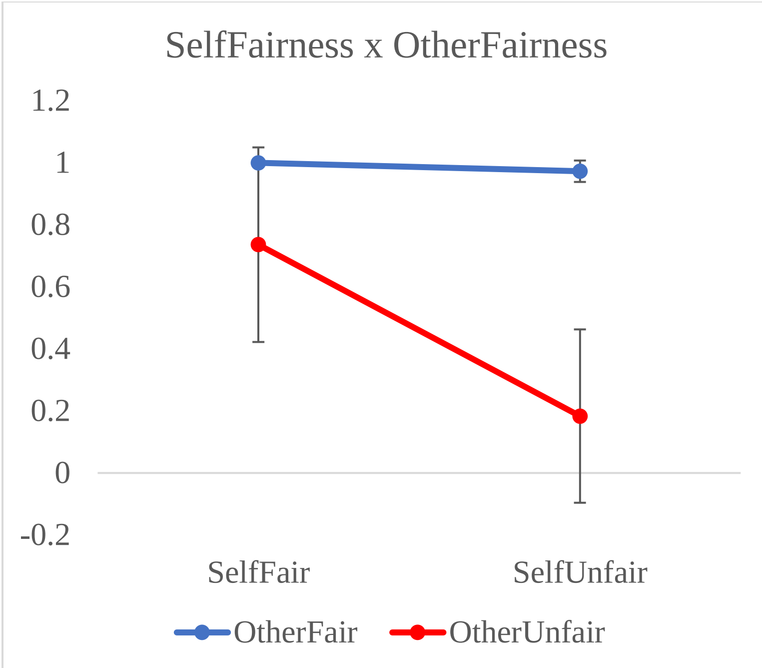


Figure 19. Study 6 SelfFairness x OtherFairness Interaction
Mean acceptance rates are shown for visualization purposes. Error bars represent SEM for mean acceptance rates.

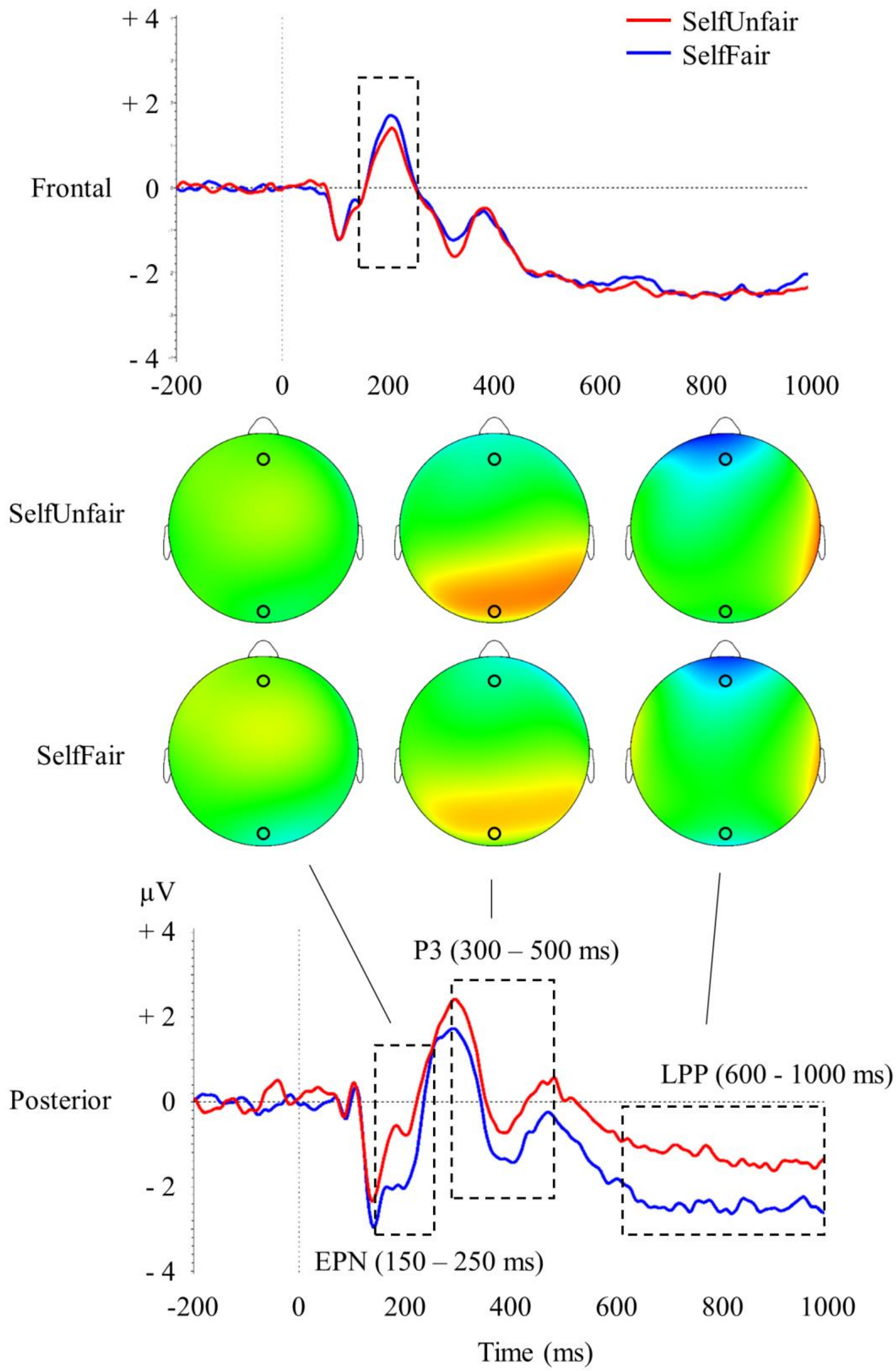


Figure 20. ERP traces and scalp plots for SelfFair and SelfUnfair distributions

Chapter 6: General Discussion

This dissertation was conducted to address two open questions: whether situational cues for self-focused or other-focused processing could influence behavior in a similar fashion to justice dispositions, and which neural systems were differentially responsive to other-focused and self-focused justice processing. Results here demonstrate that briefly adopting the perspective of a victim or observer can alter some behaviors on subsequent tasks. However, including additional measures of dispositional sensitivities to injustice for the self and others improves the explanatory power of models of participants' responses during tasks of moral judgment, distributional fairness, and antisocial behavior. Thus, perspective framing appears to provide the cues necessary for sensitive individuals to perceive and react to injustice, with downstream consequences on future behavior (Baumert, Rothmund, et al., 2013).

The second goal of this dissertation was to characterize which neural systems are associated with other-focused and self-focused justice processing. While there was no evidence for an association between justice perspectives and ERP measures of neural functioning, there was good evidence that other-focused and self-focused justice sensitivity are associated with largely distinct neural systems. The links with the executive control and salience networks suggest that these dispositional traits are associated with different processing styles, with other-focused processing relying on complex deliberations and self-focused processing arising from vigilance towards the "mean intentions" of others (Baumert, Rothmund, et al., 2013; Decety & Yoder, 2017; Gollwitzer et al., 2013; Yoder & Decety, 2014b).

6.1 Situational Cues of Justice

Overall, the perspective framing manipulation produced relatively small effects on behavior. However, this was not necessarily unexpected since the manipulation itself was subtle.

With the exception of Studies 1 and 2, a single framing block involved rating the moral permissibility of just five scenarios framed to negatively affect the self or another person. Yet, this perspective framing manipulation was predictive of changes in behavioral responses during three-party economic distributions (Study 3), the likelihood of engaging in antisocial behavior (Study 6), and also improved model fits across all six studies. However, there was strong evidence that the explanatory power of perspective framing involved interactions with dispositional justice sensitivity. Thus, some situational cues for self-focused or other-focused processing lead to changes in justice behaviors, but justice sensitivity captures individual differences in how sensitive people are to those cues.

Importantly, recent research indicates that while justice sensitivity dispositions are more variable in childhood and early adolescence, they stabilize in adulthood (Baumert, Rothmund, et al., 2013; Baumert & Schmitt, 2016; Bondü & Elsner, 2015; Bondü, Hannuschke, Elsner, & Gollwitzer, 2016; Schmitt et al., 2010). In other words, justice sensitivity dispositions in this dissertation reflect stable traits. I propose that these stable dispositions represent largely distinct processing modes which can be activated in response to situational cues, such as those provided by perspective framing. Once activated, a particular processing modes can persist over time, as evidenced by significant perspective framing effects on hemodynamic responses during tasks separated from the framing scenarios more than 10 minutes (e.g. Figures 17 and 18). However, shifting between these modes can happen relatively quickly, since changes in perspective framing occurred over a few minutes in Study 3. This also suggests that the extent to which self-focused or other-focused justice processing is dominant at a given time can be flexibly determined in response to contextual cues. Currently unknown is whether such shifts are relatively automatic, or more driven by intentional deliberation.

6.2 Neural Systems and Justice Perspectives

Study 5 indicates that self-oriented and other-oriented justice sensitivity are supported by largely distinct neural systems. Self-oriented sensitivity was associated with response in salience network and regions which are important for threat detection and avoidance, specifically striatum, MCC/SMA, ACC, and LC. Conversely, other-oriented justice sensitivity was related to ventral attentional networks and nodes in the executive control network. The association between self-oriented justice sensitivity and the salience network fits with theoretical characterizations of self-oriented sensitivity as a fear of exploitation coupled with an expectation that others are untrustworthy (Gollwitzer & Rothmund, 2011; Gollwitzer et al., 2013). Moreover, the relation between self-oriented justice sensitivity and neural response in ACC and amygdala (in both the TDT and Dots), further supports the notion that highly self-oriented individuals are vigilant to their own exploitation.

Conversely, other-oriented justice sensitivity is argued to reflect genuine concern with justice as a moral principle (Baumert, Rothmund, et al., 2013; Decety & Yoder, 2017; Schmitt et al., 2010). The results of Study 5 contribute to a small but growing body of work (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Robertson et al., 2007; Yoder & Decety, 2014b) that applying abstract principles of justice relies on complex reasoning. Taken together, these findings provide further support for the notion that self-oriented and other-oriented processing represent largely distinct processing modes, the first utilizing nodes of the salience network and the second primarily supported by the executive control network.

When viewing HighReward trials during the Dots task, other-oriented justice sensitivity was associated with increased hemodynamic response in bilateral IFG, IPL, and TPJ (Figure 13), including lateral occipital cortex (MT+) and a small region in right frontal eye fields. All of these regions have been previously identified as belonging to the ventral attention network (Corbetta & Shulman, 2002; Fox et al., 2006). From an attentional standpoint, these effects would suggest that other-oriented justice sensitivity increased the extent to which the more valuable (and rarer) stimuli were encoded as motivationally relevant (Vossel et al., 2014). However, a more recent account of IFG function argues that the primary significance of activity in vlPFC is the generation of a widespread and temporary stop signal (Sebastian et al., 2016). In this conception, activity observed within IFG is better explained as a brake on processing and behaviors in response to the detection of new information (Aron et al., 2014). This account is strengthened by the additional positive association between other-oriented justice sensitivity and activity within bilateral striatum. In contrast, dispositional self-oriented justice sensitivity was associated with greater hemodynamic response to LowReward trials within MCC, a small cluster in left TPJ, and bilaterally in hippocampus and PIC (Figures 15 and 16). These regions are reliably activated during negative experiences, including pain (Shackman et al., 2011), as well as interoception and somatosensory processing (Beissner et al., 2013). Thus, whereas other-oriented justice sensitivity was associated with reorienting to novel stimuli, self-oriented justice sensitivity instead predicted changes in interoception and potentially even frustration (Denson et al., 2008).

During the TDT, where fairness for the self was manipulated independently from fairness of the observer, self-oriented justice sensitivity was associated with hemodynamic changes in multiple regions implicated in avoiding perceived threats. These including bilateral striatum, LC, and SMA (Figure 12). A central aspect of self-oriented justice sensitivity is an expectation that others are not trustworthy, including a motivation to avoid that exploitation (Gollwitzer & Rothmund, 2011; Gollwitzer et al., 2013, 2015), as well as negative reactions to injustice (Baumert & Schmitt, 2016; Schmitt & Mohiyeddini, 1996). Based on the data from Study 5, it appears that this may result in individuals with higher self-oriented sensitivity responding to self-directed unfairness as threatening, and further suggests that these individuals may also show an accompanying change in physiological state.

In this dissertation, I have endeavored to bring together insights from psychology, behavioral economics, neuroscience, and comparative biology to investigate justice sensitivity. However, the results presented here can also inform new theories across these diverse fields of study. Recent animal work indicates that only those animals which regularly cooperate with non-kin show negative reactions to inequitable rewards (Brosnan, 2013; Brosnan & Bshary, 2016; Talbot, Price, & Brosnan, 2016). This suggests that sensitivity to inequity may have evolved to stabilize mutually beneficial behaviors in species which require regular interactions with genetically unrelated individuals (Brosnan & Bshary, 2016). From a mutualistic perspective, partner choice exerts strong pressure to at least appear to other potential interaction partners as if one were likely to reciprocate and cooperate (Baumard & Sheskin, 2015). In genetically heterogeneous populations, the most efficient proximate mechanism to establish a reputation as a reliably cooperative partner is to possess a “genuine moral sense” (Baumard et al., 2013), and justice sensitivity could represent one of these proximate mechanisms.

6.3 Future Directions

There are several important issues for future studies to investigate. First, in depth debriefing of participants or obtaining subjective affective ratings could provide some clarity into how participants made decisions during the tasks. For instance, if other-oriented justice sensitivity is related to complex reasoning about justice principles, individuals with higher trait other-oriented sensitivity would be expected to provide more complicated or elaborate answers when asked to introspect on their decision-making process. Similarly, if self-oriented sensitivity involves vigilance towards exploitation, individuals with higher self-oriented justice sensitivity should report more negative emotions in response to unfair offers. Moreover, participants were not explicitly told to attend to the identity of the target during the framing scenarios. It may be that only some participants were aware of changing target and that the effects observed here are enhanced or suppressed by recognition of the manipulation.

One limitation of this dissertation was that it combined proscriptive and prescriptive moral scenarios during perspective framing blocks. If approach and avoidance motivation play different roles in self-focused and other-focused processing, it will be crucial for future studies to probe perspective framing effects separately for proscriptive and prescriptive scenarios. The neural networks implicated by self-focused compared to other-focused processing suggest that this is likely the case.

Another potential limitation of this dissertation is that antisocial behavior was always investigated after distributional fairness. Thus, the observed associations between perspective framing or justice sensitivity and antisocial behavior could represent carryover effects. Future studies could probe this issue by counterbalancing task order or testing distributional fairness and antisocial behavior in separate samples or during separate sessions on different days.

Future studies could also employ finer gradations of unfairness for the self and others by developing a range of distributions. A broader range of distributional allocations would allow researchers to examine parametric influences of allocation amounts to the self and an observer and might also reveal different sensitivity curves for self or other fairness. Moreover, recent work (along with the results of Study 4) suggests that equality may not be the primary metric individuals use to make decisions about distributional fairness (Starmans et al., 2017). In order to tease apart preferences for equitable and equal distributions, studies could include more contextual information about the other two players, such as differences in effort, merit, and need. A full account of justice sensitivity will also require researchers to investigate neural responses when participants play the TDT in the role of Observer or Proposer.

Another important and largely unexplored area of justice research is the role of the autonomic nervous system. Such investigations may prove to be important because core nodes of the central autonomic network were implicated in self-oriented justice sensitivity, including middle cingulate and insular cortices. These regions play important roles in interoception and the coordination of homeostatic and allostatic functioning, and so precise investigation into parasympathetic and sympathetic activity could clarify differences and similarities between self-oriented and other-oriented perspectives (Craig, 2002, 2003, 2009; Critchley, Wiens, Rotshtein, Ohman, & Dolan, 2004; Medford & Critchley, 2010; Sterling & Eyer, 1988; Thayer & Lane, 2009). This line of research seems especially promising given the associations between self-oriented justice sensitivity and hemodynamic response in LC, MCC/SMA, and striatum.

6.4 Conclusions

Humans care deeply about justice, and take others' interests into account when making decisions. The studies described here find some evidence that briefly adopting the perspective of a victim to or observer of a sociomoral transgression changes this decision calculus (Study 3). However, perspective framing seems to interact with individuals' dispositional sensitivity to justice (Studies 1 and 5). In other words, self-focused and other-focused cues may interact with dispositional justice sensitivity to lead sensitive individuals to adopt different processing mode in response to contextual cues, each with downstream behavioral consequences. Moreover, dispositional self-oriented and other-oriented justice sensitivity predicted distinct activity within the salience and executive control networks, respectively, suggesting that these traits may be distinguished by vigilance towards exploitation or complex reasoning. These results provide some promising first steps towards identifying neural networks which play a critical role in justice sensitivity.

References

Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *The Neuroscientist*, *17*(5), 560–574. http://doi.org/10.1177/1073858410393359

Akaike, H. (1969). Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, *22*, 203–217.

Alexopoulos, J., Pfabigan, D. M., Göschl, F., Bauer, H., & Fischmeister, F. P. S. (2013). Agency matters! Social preferences in the three-person ultimatum game. *Frontiers in Human Neuroscience*, *7*(312), 1–10. http://doi.org/10.3389/fnhum.2013.00312

Alexopoulos, J., Pfabigan, D. M., Lamm, C., Bauer, H., & Fischmeister, F. P. S. (2012). Do we care about the powerless third? An ERP study of the three-person ultimatum game. *Frontiers in Human Neuroscience*, *6*(59), 1–9. http://doi.org/10.3389/fnhum.2012.00059

Almås, I., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*(5982), 1176–1178.

Apps, M. A. J., Rushworth, M. F. S., & Chang, S. W. C. (2016). The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron*, *90*(4), 692–707. http://doi.org/10.1016/j.neuron.2016.04.018

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*, *18*(4), 177–185. http://doi.org/10.1016/j.tics.2013.12.003

Avram, M., Hennig-Fast, K., Bao, Y., Pöppel, E., Reiser, M., Blautzik, J., … Gutyrchik, E. (2014). Neural correlates of moral judgments in first- and third-person perspectives: implications for neuroethics and beyond. *BMC Neuroscience*, *15*, 39.

http://doi.org/10.1186/1471-2202-15-39

Báez-Mendoza, R., Harris, C. J., & Schultz, W. (2013). Activity of striatal neurons reflects social action and own reward. *Proceedings of the National Academy of Sciences*, *110*(41), 16634–16639. http://doi.org/10.1073/pnas.1211342110

Báez-Mendoza, R., van Coeverden, C. R., & Schultz, W. (2016). A neuronal reward inequity signal in primate striatum. *Journal of Neurophysiology*, *115*(1), 68–79. http://doi.org/10.1152/jn.00321.2015

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends in Neurosciences*, *29*(5), 272–9. http://doi.org/10.1016/j.tins.2006.03.002

Bamford, S., Broyd, S. J., Benikos, N., Ward, R., Wiersema, J. R., & Sonuga-barke, E. (2015). The late positive potential: A neural marker of the regulation of emotion-based approach-avoidance actions? *Biological Psychology*, *105*, 1–9. http://doi.org/10.1016/j.biopsycho.2015.01.009

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 51. http://doi.org/10.18637/jss.v067.i01

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *The Social Neuroscience of Empathy* (pp. 3–15). Cambridge, MA: MIT Press.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, *36*(1), 59–78. http://doi.org/10.1017/S0140525X11002202

Baumard, N., & Sheskin, M. (2015). Partner choice and the evolution of a contractualist morality. In J. Decety & T. Wheatley (Eds.), *The Moral Brain* (pp. 35–48). Cambridge, MA: MIT Press.

Baumert, A., Beierlein, C., Schmitt, M., Kemper, C. J., Kovaleva, A., Liebig, S., & Rammstedt, B. (2014). Measuring four perspectives of justice sensitivity with two items each. *Journal of Personality Assessment*, *96*(3), 380–90. http://doi.org/10.1080/00223891.2013.836526

Baumert, A., Gollwitzer, M., Staubach, M., & Schmitt, M. (2011). Justice sensitivity and the processing of justice-related information. *European Journal of Personality*, *25*, 386–397. http://doi.org/10.1002/per

Baumert, A., Halmburger, A., & Schmitt, M. (2013a). Interventions against norm violations: Dispositional determinants of self-reported and real moral courage. *Personality and Social Psychology Bulletin*, *39*(8), 1053–1068. http://doi.org/10.1177/0146167213490032

Baumert, A., Halmburger, A., & Schmitt, M. (2013b). Interventions against norm violations: Dispositional determinants of self-reported and real moral courage. *Personality and Social Psychology Bulletin*, *39*(8), 1053–1068. http://doi.org/10.1177/0146167213490032

Baumert, A., Hangarter, M., Gollwitzer, M., & Schmitt, M. (2007). *Attentional processes guided by justice sensitivity: Evidence from a modified Stroop task. Berichte aus der Arbeitsgruppe "'Verantwortung, Gerechtigkeit, Moral.'"* Trier, Germany.

Baumert, A., Otto, K., Thomas, N., Bobocel, D. R., & Schmitt, M. (2012). Processing of unjust and just information: interpretation and memory performance related to dispositional victim sensitivity. *European Journal of Personality*, *26*, 99–110. http://doi.org/10.1002/per

Baumert, A., Rothmund, T., Thomas, N., Gollwitzer, M., & Schmitt, M. J. (2013). Justice as a moral motive: Belief in a just world and justice sensitivity as potential indicator of the justice motive. In K. Heinrichs, F. Oser, & T. Lovat (Eds.), *Handbook of Moral Motivation* (pp. 159–180). Rotterdam: Sense Publishers.

Baumert, A., Schlösser, T., & Schmitt, M. (2014). Economic games: A performance-based assessment of fairness and altruism. *European Journal of Psychological Assessment*, *30*(3), 178–192. http://doi.org/10.1027/1015-5759/a000183

Baumert, A., & Schmitt, M. (2009). Justice-sensitive interpretations of ambiguous situations. *Australian Journal of Psychology*, *61*(1), 6–12. http://doi.org/10.1080/00049530802607597

Baumert, A., & Schmitt, M. (2016). Justice Sensitivity. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of Social Justice Theory and Research* (pp. 161–180). New York, NY: Springer. http://doi.org/10.1007/978-1-4939-3216-0

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, *64*(5), 756–770. http://doi.org/10.1016/j.neuron.2009.11.017

Beissner, F., Meissner, K., Bär, K.-J., & Napadow, V. (2013). The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. *The Journal of Neuroscience*, *33*(25), 10503–11. http://doi.org/10.1523/JNEUROSCI.1103-13.2013

Benajmini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Benarroch, E. E. (1993). The central autonomic network: Functional organization, dysfunction, and perspective. *Mayo Clinic Proceedings*, *68*(10), 988–1001. http://doi.org/10.1016/S0025-6196(12)62272-1

Berridge, C. W. (2008). Noradrenergic modulation of arousal. *Brain Research Reviews*, *58*(1), 1–17. http://doi.org/10.1016/j.brainresrev.2007.10.013

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, *42*(1), 33–84. http://doi.org/10.1016/S0165-0173(03)00143-7

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure systems in the brain. *Neuron*, *86*(3), 646–664. http://doi.org/10.1016/j.neuron.2015.02.018

Boksem, M. A. S., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, *5*(1), 118–28. http://doi.org/10.1080/17470910903202666

Bolton, G. E., & Ockenfels, A. (2000). ERC : A theory of equity, reciprocity, and competition. *The American Economic Review*, *90*(1), 166–193.

Bondü, R., & Elsner, B. (2015). Justice sensitivity in childhood and adolescence. *Social Development*, *24*(2), 420–441. http://doi.org/10.1111/sode.12098

Bondü, R., Hannuschke, M., Elsner, B., & Gollwitzer, M. (2016). Inter-individual stabilization of justice sensitivity in childhood and adolescence. *Journal of Research in Personality*, *64*, 11–20. http://doi.org/10.1016/j.jrp.2016.06.021

Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An FMRI investigation.

Journal of Cognitive Neuroscience, *18*(5), 803–17. http://doi.org/10.1162/jocn.2006.18.5.803

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–46. http://doi.org/10.1016/j.tics.2004.10.003

Brosnan, S. F. (2013). Justice- and fairness-related behaviors in nonhuman primates. *Proceedings of the National Academy of Sciences*, *110*, 10416–10423. http://doi.org/10.1073/pnas.1301194110

Brosnan, S. F., & Bshary, R. (2016). On potential links between inequity aversion and the structure of interactions for the evolution of cooperation. *Behaviour*, *153*(9–11), 1267–1292. http://doi.org/10.1163/1568539X-00003355

Buckholtz, J. W., Asplund, C. L., Dux, P. E., Zald, D. H., Gore, J. C., Jones, O. D., & Marois, R. (2008). The neural correlates of third-party punishment. *Neuron*, *60*(5), 930–40. http://doi.org/10.1016/j.neuron.2008.10.016

Buckholtz, J. W., & Marois, R. (2012). The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience*, *15*(5), 655–61. http://doi.org/10.1038/nn.3087

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. http://doi.org/10.1196/annals.1440.011

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet kigh-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5.

http://doi.org/10.1177/1745691610393980

Bzdok, D., Laird, A. R., Zilles, K., Fox, P. T., & Eickhoff, S. B. (2013). An investigation of the structural, connectional, and functional subspecialization in the human amygdala. *Human Brain Mapping*, *34*(12), 3247–66. http://doi.org/10.1002/hbm.22138

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure & Function*, *217*(4), 783–96. http://doi.org/10.1007/s00429-012-0380-y

Cacioppo, J. T., Crites, S. L., & Gardner, W. L. (1996). Attitudes to the right: Evaluative processing is associated with lateralized late positive event-related brain potentials. *Personality and Social Psychology Bulletin*, *22*(12), 1205–1219. http://doi.org/10.1177/01461672962212002

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, *76*(5), 839–855. http://doi.org/10.1037/0022-3514.76.5.839

Cappelen, A. W., Eichele, T., Hugdahl, K., Specht, K., Sorensen, E. O., & Tungodden, B. (2014). Equity theory and fair inequality: A neuroeconomic study. *Proceedings of the National Academy of Sciences*, *111*(43), 15368–15372. http://doi.org/10.1073/pnas.1414602111

Carnes, N. C., Lickel, B., & Janoff-Bulman, R. (2015). Shared perceptions: Morality is embedded in social contexts. *Personality and Social Psychology Bulletin*, *41*(3), 351–362. http://doi.org/10.1177/0146167214566187

Carrión, R. E., Keenan, J. P., & Sebanz, N. (2010). A truth that's told with bad intent: An ERP study of deception. *Cognition*, *114*(1), 105–110. http://doi.org/10.1016/j.cognition.2009.05.014

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. http://doi.org/10.1037/0022-3514.67.2.319

Cheng, Y., Hung, A.-Y., & Decety, J. (2012). Dissociation between affective sharing and emotion understanding in juvenile psychopaths. *Development and Psychopathology*, *24*(2), 623–36. http://doi.org/10.1017/S095457941200020X

Christensen, R. H. B. (2015). ordinal - Regression models for ordinal data.

Civai, C. (2013). Rejecting unfairness: Emotion-driven reaction or cognitive heuristic? *Frontiers in Human Neuroscience*, *7*, 126. http://doi.org/10.3389/fnhum.2013.00126

Civai, C., Corradi-Dell'Acqua, C., Gamer, M., & Rumiati, R. I. (2010). Are irrational reactions to unfairness truly emotionally-driven? Dissociated behavioural and emotional responses in the Ultimatum Game task. *Cognition*, *114*(1), 89–95. http://doi.org/10.1016/j.cognition.2009.09.001

Civai, C., Crescentini, C., Rustichini, A., & Rumiati, R. I. (2012). Equality versus self-interest in the brain: Differential roles of anterior insula and medial prefrontal cortex. *NeuroImage*, *62*(1), 102–112. http://doi.org/10.1016/j.neuroimage.2012.04.037

Civai, C., Miniussi, C., & Rumiati, R. I. (2015). Medial prefrontal cortex reacts to unfairness if this damages the self: A tDCS study. *Social Cognitive and Affective Neuroscience*, *10*, 1054–1060. http://doi.org/10.1093/scan/nsu154

Cohen, J. D., Botvinick, M. M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nature Neuroscience*, *3*(5), 421–3. http://doi.org/10.1038/74783

Cooper, A., & Gomez, R. (2008). The development of a short form of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire. *Journal of Individual Differences*, *29*(2), 90–104. http://doi.org/10.1027/1614-0001.29.2.90

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–15. http://doi.org/10.1038/nrn755

Corradi-Dell'Acqua, C., Civai, C., Rumiati, R. I., & Fink, G. R. (2013). Disentangling self- and fairness-related neural mechanisms involved in the ultimatum game: An fMRI study. *Social Cognitive and Affective Neuroscience*, *8*(4), 424–431. http://doi.org/10.1093/scan/nss014

Cowell, J. M., & Decety, J. (2015a). Precursors to morality in development as a complex interplay between neural, socioenvironmental, and behavioral facets. *Proceedings of the National Academy of Sciences*, *112*(41), 12657–12662. http://doi.org/10.1073/pnas.1508832112

Cowell, J. M., & Decety, J. (2015b). The neuroscience of implicit moral evaluation and its relation to generosity in early childhood. *Current Biology*, *25*(1), 93–97. http://doi.org/10.1016/j.cub.2014.11.002

Cowell, J. M., Lee, K., Malcolm-Smith, S., Selcuk, B., Zhou, X., & Decety, J. (2016). The development of generosity and moral cognition across five cultures. *Developmental Science*, 1–12. http://doi.org/10.1111/desc.12403

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews. Neuroscience*, *3*(8), 655–66. http://doi.org/10.1038/nrn894

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology*, *13*(4), 500–505. http://doi.org/10.1016/S0959-4388(03)00090-4

Craig, A. D. (2009). How do you feel--now? The anterior insula and human awareness. *Nature Reviews. Neuroscience*, *10*(1), 59–70. http://doi.org/10.1038/nrn2555

Crawford, L. E., & Cacioppo, J. T. (2002). Learning where to look for danger: Integrating affective and spatial information. *Psychological Science*, *13*(5), 449–53. http://doi.org/10.1111/1467-9280.00479

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–95. http://doi.org/10.1038/nn1176

Crockett, M. J., Clark, L., Lieberman, M. D., Tabibnia, G., & Robbins, T. W. (2010). Impulsive choice and altruistic punishment are correlated and increase in tandem with serotonin depletion. *Emotion*, *10*(6), 855–862. http://doi.org/10.1037/a0019861

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. http://doi.org/10.1016/j.jesp.2015.07.006

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111.

Dar, Y., & Resh, N. (2001). Exploring the multifaceted structure of sense of deprivation. *European Journal of Social Psychology*, *31*(1), 63–81. http://doi.org/10.1002/ejsp.32

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. http://doi.org/10.1037//0022-3514.44.1.113

Davis, M., & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry*, *6*(1), 13–34.

Dawes, C. T., Loewen, P. J., Schreiber, D., Simmons, A. N., Flagan, T., McElreath, R., … Paulus, M. P. (2012). Neural basis of egalitarian behavior. *Proceedings of the National Academy of Sciences*, *109*(17), 6479–6483. http://doi.org/10.1073/pnas.1118653109

Decety, J. (2015). The neural pathways, development and functions of empathy. *Current Opinion in Behavioral Sciences*, *3*, 1–6. http://doi.org/10.1016/j.cobeha.2014.12.001

Decety, J., & Cowell, J. M. (2014). The complex relation between morality and empathy. *Trends in Cognitive Sciences*, *18*(7), 337–339. http://doi.org/10.1016/j.tics.2014.04.008

Decety, J., & Cowell, J. M. (2015). Empathy, justice, and moral behavior. *AJOB Neuroscience*, *6*(3), 3–14. http://doi.org/10.1080/21507740.2015.1047055

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*(2), 71–100. http://doi.org/10.1177/1534582304267187

Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: A neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–20.

http://doi.org/10.1093/cercor/bhr111

Decety, J., & Wheatley, T. (2015). *The Moral Brain: A Multidisciplinary Perspective*. Cambridge, MA: MIT Press.

Decety, J., Yang, C.-Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: an event-related brain potential study. *NeuroImage*, *50*(4), 1676–82. http://doi.org/10.1016/j.neuroimage.2010.01.025

Decety, J., & Yoder, K. J. (2015). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social Neuroscience*, *11*(1), 1–14. http://doi.org/10.1080/17470919.2015.1029593

Decety, J., & Yoder, K. J. (2017). The emerging social neuroscience of justice motivation. *Trends in Cognitive Sciences*, *21*(1), 6–14. http://doi.org/10.1016/j.tics.2016.10.008

Denson, T. F., Pedersen, W. C., Ronquillo, J., & Nandy, A. S. (2008). The angry brain: Neural correlates of anger, angry rumination, and aggressive personality. *Journal of Cognitive Neuroscience*, *21*(4), 734–744. http://doi.org/10.1162/jocn.2009.21051

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology*, *35*(1), 67–82. http://doi.org/10.1016/j.psyneuen.2009.10.006

Dorsey, D. (2013). The supererogatory, and how to accommodate it. *Utilitas*, *25*(3), 355–382. http://doi.org/10.1017/S095382081200060X

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A.

T., … Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, *104*(26), 11073–8. http://doi.org/10.1073/pnas.0704320104

Edele, A., Dziobek, I., & Keller, M. (2013). Explaining altruistic sharing in the dictator game: The role of affective empathy, cognitive empathy, and justice sensitivity. *Learning and Individual Differences*, *24*, 96–102. http://doi.org/10.1016/j.lindif.2012.12.020

Eldar, E., Hauser, T. U., Dayan, P., & Dolan, R. J. (2016). Striatal structure and function predict individual biases in learning to avoid pain. *Proceedings of the National Academy of Sciences*, *113*(17), 4812–4817. http://doi.org/10.1073/pnas.1519829113

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, *46*(1), 160–73. http://doi.org/10.1016/j.neuropsychologia.2007.07.023

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences*, *11*(10), 419–427. http://doi.org/10.1016/j.tics.2007.09.002

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, *8*(4), 185–190. http://doi.org/10.1016/j.tics.2004.02.007

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868. http://doi.org/10.1162/003355399556151

Feng, C., Luo, Y. J., & Krueger, F. (2015). Neural signatures of fairness-related normative decision making in the ultimatum game: A coordinate-based meta-analysis. *Human Brain Mapping*, *36*(2), 591–602. http://doi.org/10.1002/hbm.22649

Ferry, M. (2013). Does morality demand our very best? On moral prescriptions and the line of duty. *Philosophical Studies*, *165*(2), 573–589. http://doi.org/10.1007/s11098-012-9968-6

Fetchenhauer, D., & Huang, X. (2004). Justice sensitivity and distributive decisions in experimental games. *Personality and Individual Differences*, *36*(5), 1015–1029. http://doi.org/10.1016/S0191-8869(03)00197-1

Foa, U. G., & Foa, E. B. (1974). *Societal Structures of the Mind*. Springfield, IL: Charles C. Thomas.

Folger, R. (2011). Fairness as deonance. In S. Gilliland, D. Steiner, & D. Skarlicki (Eds.), *Theoretical and Cultural Perspectives on Organizational Justice* (pp. 3–33). Grennwich, CT: Information Age Publishing.

Folstein, J., & Petten, C. Van. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*(1), 152–170.

Foot, P. (1967). *The problem of abortion and the doctrine of double effect*. Oxford, UK: Blackwell.

Fox, M. D., Corbetta, M., Synder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, *103*(26), 10046–10051.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678. http://doi.org/10.1073/pnas.0504136102

Gabay, A. S., Radua, J., Kempton, M. J., & Mehta, M. A. (2014). The Ultimatum Game and the brain: A meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *47*, 549–558. http://doi.org/10.1016/j.neubiorev.2014.10.014

Garrett, N., Lazzaro, S. C., Ariely, D., & Sharot, T. (2016). The brain adapts to dishonesty. *Nature Neuroscience*, *19*(October), 1–6. http://doi.org/10.1038/nn.4426

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–82. http://doi.org/10.1126/science.1066893

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Gino, F. (2015). Understanding ordinary unethical behavior: Why people who value morality act immorally. *Current Opinion in Behavioral Sciences*, *3*(im), 107–111. http://doi.org/10.1016/j.cobeha.2015.03.001

Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, *102*(3), 445–459. http://doi.org/10.1037/a0026406

Glenn, A. L., & Raine, A. (2014). Neurocriminology: Implications for the punishment, prediction and prevention of criminal behaviour. *Nature Publishing Group*, *15*(1), 54–63. http://doi.org/10.1038/nrn3640

Gollwitzer, M., & Rothmund, T. (2011). What exactly are victim-sensitive persons sensitive to? *Journal of Research in Personality*, *45*(5), 448–455. http://doi.org/10.1016/j.jrp.2011.05.003

Gollwitzer, M., Rothmund, T., Alt, B., & Jekel, M. (2012). Victim sensitivity and the accuracy of social judgments. *Personality and Social Psychology Bulletin*, *38*(8), 975–984. http://doi.org/10.1177/0146167212440887

Gollwitzer, M., Rothmund, T., Pfeiffer, A., & Ensenbach, C. (2009). Why and when justice sensitivity leads to pro- and antisocial behavior. *Journal of Research in Personality*, *43*(6), 999–1005. http://doi.org/10.1016/j.jrp.2009.07.003

Gollwitzer, M., Rothmund, T., & Süssenbach, P. (2013). The Sensitivity to Mean Intentions (SeMI) model: Basic assumptions, recent findings, and potential avenues for future research. *Social and Personality Psychology Compass*, *7*(7), 415–426. http://doi.org/10.1111/spc3.12041

Gollwitzer, M., Schmitt, M. J., Schalke, R., Maes, J., & Baer, A. (2005). Asymmetrical effects of justice sensitivity perspectives on prosocial and antisocial behavior. *Social Justice Research*, *18*(2), 183–201. http://doi.org/10.1007/s11211-005-7368-1

Gollwitzer, M., Süssenbach, P., & Hannuschke, M. (2015). Victimization experiences and the stabilization of victim sensitivity. *Frontiers in Psychology*, *6*, 439. http://doi.org/10.3389/fpsyg.2015.00439

Gospic, K., Mohlin, E., Fransson, P., Petrovic, P., Johannesson, M., & Ingvar, M. (2011). Limbic justice—Amygdala involvement in immediate rejection in the Ultimatum Game. *PLoS Biology*, *9*(5), e1001054. http://doi.org/10.1371/journal.pbio.1001054

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–85. http://doi.org/10.1037/a0021847

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion*, *4*(3), 269–288.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144–54. http://doi.org/10.1016/j.cognition.2007.11.004

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400. http://doi.org/10.1016/j.neuron.2004.09.027

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences*, *106*(30), 12506–12511. http://doi.org/10.1073/pnas.0900152106

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–8. http://doi.org/10.1126/science.1062872

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4), 367–388. http://doi.org/10.1016/0167-2681(82)90011-7

Hafer, C. L., & Begue, L. (2005). Experimental research on just-world theory: Problems, developments, and future challenges. *Psychological Bulletin*, *131*(1), 128–167. http://doi.org/10.1037/0033-2909.131.1.128

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–70.

http://doi.org/10.1111/j.1469-8986.2005.00278.x

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology*, *35*(2), 129–55. http://doi.org/10.1080/87565640903526504

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, *42*(2), 151–60. http://doi.org/10.1111/j.1469-8986.2005.00270.x

Hamlin, J. K. (2014). The origins of human morality: Complex socio-moral evaluations by preverbal infants. In J. Decety & Y. Christen (Eds.), *New Frontiers in Social Neuroscience* (pp. 165–188). New York, NY: Springer.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience*, *28*(22), 5623–5630. http://doi.org/10.1523/JNEUROSCI.1309-08.2008

Haroush, K., & Williams, Z. M. (2015). Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell*, *160*(6), 1233–1245. http://doi.org/10.1016/j.cell.2015.01.045

Harrison, L. A., Hurlemann, R., & Adolphs, R. (2015). An enhanced default approach bias following amygdala lesions in humans. *Psychological Science*. http://doi.org/10.1177/0956797615583804

Harsay, H. A., Spaan, M., Wijnen, J. G., & Ridderinkhof, K. R. (2012). Error awareness and salience processing in the oddball task: Shared neural mechanisms. *Frontiers in Human Neuroscience*, *6*(August), 246. http://doi.org/10.3389/fnhum.2012.00246

Henrich, J., Ensminger, J., Mcelreath, R., Barr, A., Barrett, C., Bolyanatz, A., … Gwako, E. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*, 1480–1484. http://doi.org/10.1126/science.1182238

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., … Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767–1770. http://doi.org/10.1126/science.1127333

Higgins, E. T. (1997). Beyond pleasure and pain. *The American Psychologist*, *52*(12), 1280–1300. http://doi.org/10.1037/0003-066X.52.12.1280

Hosokawa, T., & Watanabe, M. (2015). Egalitarian reward contingency in competitive games and primate prefrontal neuronal activity. *Frontiers in Neuroscience*, *9*, 165. http://doi.org/10.3389/fnins.2015.00165

Hu, X., Wu, H., & Fu, G. (2011). Temporal course of executive control when lying about self- and other-referential information: An ERP study. *Brain Research*, *1369*, 149–157. http://doi.org/10.1016/j.brainres.2010.10.106

Hutzler, F. (2014). Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *NeuroImage*, *84*, 1061–1069. http://doi.org/10.1016/j.neuroimage.2012.12.075

Ito, A., Abe, N., Fujii, T., Hayashi, A., Ueno, A., Mugikura, S., … Mori, E. (2012). The contribution of the dorsolateral prefrontal cortex to the preparation for deception and truth-telling. *Brain Research*, *1464*, 43–52. http://doi.org/10.1016/j.brainres.2012.05.004

Ito, A., Abe, N., Fujii, T., Ueno, A., Koseki, Y., Hashimoto, R., … Mori, E. (2011). The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. *Neuroscience Research*, *69*(2), 121–128. http://doi.org/10.1016/j.neures.2010.11.001

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape: Moral motives and group-based moralities. *Personality and Social Psychology Review*, *17*(3), 219–236. http://doi.org/10.1177/1088868313480274

Janoff-Bulman, R., Sheikh, S., & Baldacci, K. G. (2008). Mapping moral motives: Approach, avoidance, and political orientation. *Journal of Experimental Social Psychology*, *44*(4), 1091–1099. http://doi.org/10.1016/j.jesp.2007.11.003

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: two faces of moral regulation. *Journal of Personality and Social Psychology*, *96*(3), 521–537. http://doi.org/10.1037/a0013779

Jensen, K., Vaish, A., & Schmidt, M. F. H. (2014). The emergence of human prosociality: Aligning with others through feelings, concerns, and norms. *Frontiers in Psychology*, *5*, 822. http://doi.org/10.3389/fpsyg.2014.00822

Kahane, G. (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. *Social Neuroscience*, *10*(5), 551–560. http://doi.org/10.1080/17470919.2015.1023400

Kahane, G., Everett, J. A. C., Earp, B. D., Farias, M., & Savulescu, J. (2015). “Utilitarian” judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193–209. http://doi.org/10.1016/j.cognition.2014.10.005

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. *PLoS ONE*, *7*(8), e43979. http://doi.org/10.1371/journal.pone.0043979

Kayser, D. N., Greitemeyer, T., Fischer, P., & Frey, D. (2010). Why mood affects help giving, but not moral courage: Comparing two types of prosocial behaviour. *European Journal of Social Psychology*, *40*, 1136–1157. http://doi.org/10.1002/ejsp

Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, *39*(5), 641–9. http://doi.org/10.1017.S0048577202394162

Koban, L., Pourtois, G., Bediou, B., & Vuilleumier, P. (2012). Effects of social context and predictive relevance on action outcome monitoring. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 460–78. http://doi.org/10.3758/s13415-012-0091-0

Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, *7*(6), 708–14. http://doi.org/10.1093/scan/nsr048

Koenigs, M., & Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the Ultimatum Game. *Journal of Neuroscience*, *27*(4), 951–956. http://doi.org/10.1523/JNEUROSCI.4606-06.2007

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*(7138), 908–11. http://doi.org/10.1038/nature05631

Kohlberg, L. (1968). The child as a moral philosopher. *Psychology Today*, *2*, 25–30.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in linear mixed effects models.

Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences*, *111*(6), 2343–2348. http://doi.org/10.1073/pnas.1321596111

Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: reflex foundations in animal and human research. *Progress in Brain Research*, *156*, 3–29. http://doi.org/10.1016/S0079-6123(06)56001-7

Langleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*(4), 262–272. http://doi.org/10.1002/hbm.20191

Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., … Childress, A. R. (2002). Brain Activity during Simulated Deception: An Event-Related Functional Magnetic Resonance Study. *NeuroImage*, *15*(3), 727–732. http://doi.org/10.1006/nimg.2001.1003

Lee, T. M. C., Lee, T. M. Y., Raine, A., & Chan, C. C. H. (2010). Lying about the valence of affective pictures: An fMRI study. *PLoS ONE*, *5*(8). http://doi.org/10.1371/journal.pone.0012291

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. http://doi.org/10.18637/jss.v069.i01

Lerner, M. J. (1980). *The belief in a justice world: A fundamental delusion*. New York, NY: Plenum Press.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038.

http://doi.org/10.1016/j.conb.2012.06.001

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social Cognitive and Affective Neuroscience*, *4*(4), 423–8. http://doi.org/10.1093/scan/nsp052

Lotz, S., Baumert, A., Schlösser, T., Gresser, F., & Fetchenhauer, D. (2011). Individual differences in third-party interventions: How justice sensitivity shapes altruistic punishment. *Negotiation and Conflict Management Research*, *4*(4), 297–313.

Lotz, S., Schlösser, T., Cain, D. M., & Fetchenhauer, D. (2013). The (in)stability of social preferences: Using justice sensitivity to predict when altruism collapses. *Journal of Economic Behavior and Organization*, *93*, 141–148. http://doi.org/10.1016/j.jebo.2013.07.012

Lovas, L., & Wolt, R. (2002). Sensitivity to injustice in the context of some personality traits. *Studia Psychologica*, *44*(2), 125–131.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed.). Cambridge, MA: MIT Press.

Luck, S. J., & Kappenman, E. S. (Eds.). (2012). *The Oxford Handbook of Event-Related Potential Components*. New York, NY: Oxford University Press.

Ma, Q., & Hu, Y. (2015). Beauty matters: Social preferences in a three-person ultimatum game. *PLoS ONE*, *10*(5), 1–17. http://doi.org/10.1371/journal.pone.0125806

Macnamara, A., Foti, D., & Hajcak, G. (2009). Tell me about it: neural activity elicited by emotional pictures and preceding descriptions. *Emotion*, *9*(4), 531–43.

http://doi.org/10.1037/a0016251

Malle, B. F., & Guglielmo, S. (2012). Are intentionality judgments fundamentally moral? In R. Langdon & C. Mackenzie (Eds.), *Emotions, imagination, and moral reasoning* (pp. 275–293). New York, NY: Psychology Press.

Massi, B., & Luhmann, C. C. (2015). Fairness influences early signatures of reward-related neural processing. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 768–775. http://doi.org/10.3758/s13415-015-0362-7

Mather, M., Cacioppo, J. T., & Kanwisher, N. (2013). How fMRI can inform cognitive theories. *Perspectives on Psychological Science*, *8*(1), 108–113. http://doi.org/10.1177/1745691612469037.How

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016). Norepinephrine ignites local hot spots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, *39*, 1–19. http://doi.org/10.1017/S0140525X15000667

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459*(7248), 837–841. http://doi.org/10.1038/nature08028

Mazaika, P. K., Hoeft, F., Glover, G. H., & Reiss, A. L. (2009). Methods and software for fMRI analysis for clinical subjects. In *Presentation at the 15th Annual Meeting of the Organization for Human Brain Mapping*.

Mazar, N., & Zhong, C.-B. (2010). Do green products make us better people? *Psychological Science*, *21*(4), 494–498. http://doi.org/10.1177/0956797610363538

Medford, N., & Critchley, H. D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Structure & Function*, *214*(5–6), 535–49. http://doi.org/10.1007/s00429-010-0265-x

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, *214*(5–6), 655–667. http://doi.org/10.1007/s00429-010-0262-0.Saliency

Miller, A. R., & Rosenfeld, J. P. (2004). Response-specific scalp distributions in deception detection and ERP correlates of psychopathic personality traits. *Journal of Psychophysiology*, *18*(1), 13–26. http://doi.org/10.1027/0269-8803.18.1.13

Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*, *14*(3), 1–39. http://doi.org/10.1037/a0035361

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, *28*(4), 757–62. http://doi.org/10.1016/j.neuroimage.2005.03.011

Mohiyeddini, C., & Schmitt, M. J. (1997). Sensitivity to befallen injustice and reactions to unfair treatment in a laboratory situation. *Social Justice Research*, *10*(3), 333–353. http://doi.org/10.1007/BF02683307

Moll, J., de Oliveira-Souza, R., Garrido, G. J., Bramati, I. E., Caparelli-Daquer, E. M. A., Paiva, M. L. M. F., … Grafman, J. (2007). The self as a moral agent: Linking the neural bases of social agency and moral sensitivity. *Social Neuroscience*, *2*(3–4), 336–52. http://doi.org/10.1080/17470910701392024

Moore, C., & Gino, F. (2013). Ethically adrift: How others pull our moral compass from true North, and how we can fix it. *Research in Organizational Behavior*, *33*, 53–77. http://doi.org/10.1016/j.riob.2013.08.001

Moran, T. P., Jendrusina, A. A., & Moser, J. S. (2013). The psychometric properties of the late positive potential during emotion processing and regulation. *Brain Research*, *1516*, 66–75. http://doi.org/10.1016/j.brainres.2013.04.018

Morrison, S. E., & Salzman, C. D. (2010). Re-valuing the amygdala. *Current Opinion in Neurobiology*, *20*(2), 221–230. http://doi.org/10.1016/j.conb.2010.02.007

Mussel, P., Hewig, J., Allen, J. J. B., Coles, M. G. H., & Miltner, W. (2014). Smiling faces, sometimes they don't tell the truth: Facial expression in the ultimatum game impacts decision making and event-related potentials. *Psychophysiology*, *51*(4), 358–363. http://doi.org/10.1111/psyp.12184

Nieuwenhuys, R. (2012). *The insular cortex: A review*. *Progress in Brain Research* (1st ed., Vol. 195). Elsevier B.V. http://doi.org/10.1016/B978-0-444-53860-4.00007-6

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development*, *49*(2), 400–407.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776. http://doi.org/10.1016/j.conb.2004.10.016

O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, *33*, 815–826. http://doi.org/10.1016/S0896-6273(02)00603-7

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting snakes in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478. http://doi.org/10.1037/AXJ96-3445.130.3.466

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological Psychology*, *77*(3), 247–65. http://doi.org/10.1016/j.biopsycho.2007.11.006

Patil, I., & Silani, G. (2014). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in Psychology*, *5*(May), 1–12. http://doi.org/10.3389/fpsyg.2014.00501

Perry, A., Bentin, S., Bartal, I. B.-A., Lamm, C., & Decety, J. (2010). "Feeling" the pain of those who are different from us: Modulation of EEG in the mu/alpha range. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(4), 493–504. http://doi.org/10.3758/CABN.10.4.493

Pessiglione, M., & Delgado, M. R. (2015). The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. *Current Opinion in Behavioral Sciences*, *5*, 78–84. http://doi.org/10.1016/j.cobeha.2015.08.006

Peterburs, J., Voegler, R., Liepelt, R., Schulze, A., Wilhelm, S., Ocklenburg, S., & Straube, T. (2017). Processing of fair and unfair offers in the ultimatum game under social observation. *Scientific Reports*, *7*(44062), 1–12. http://doi.org/10.1038/srep44062

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–48. http://doi.org/10.1016/j.clinph.2007.04.019

Qu, C., Wang, Y., & Huang, Y. (2013). Social exclusion modulates fairness consideration in the ultimatum game: An ERP study. *Frontiers in Human Neuroscience*, *7*(August), 505.

http://doi.org/10.3389/fnhum.2013.00505

Raichle, M. E. (2015). The Brain's Default Mode Network. *Annual Review of Neuroscience, 38*, 433–447. http://doi.org/10.1196/annals.1440.011

Reynolds, S. J., Dang, C. T., Yam, K. C., & Leavitt, K. (2014). The role of moral knowledge in everyday immorality: What does it matter if I know what is right? *Organizational Behavior and Human Decision Processes*, *123*(2), 124–137. http://doi.org/10.1016/j.obhdp.2013.10.008

Robertson, D., Snarey, J., Ousley, O., Harenski, K., DuBois Bowman, F., Gilkey, R., & Kilts, C. (2007). The neural processing of moral sensitivity to issues of justice and care. *Neuropsychologia*, *45*(4), 755–66. http://doi.org/10.1016/j.neuropsychologia.2006.08.014

Rosas, A., & Koenigs, M. (2014). Beyond "utilitarianism": Maximizing the clinical impact of moral judgment research. *Social Neuroscience*, *9*(6), 661–7. http://doi.org/10.1080/17470919.2014.937506

Ross, L., & Nisbett, R. E. (1991). *The Person and the Situation*. Boston, MA: McGraw-Hill.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562. http://doi.org/10.1038/nrn3776

Ruff, C. C., Ugazio, G., & Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, *342*(6157), 482–484. http://doi.org/10.1126/science.1241399

Sabbagh, C., & Schmitt, M. (Eds.). (2016). *Handbook of Social Justice Theory and Research*. New York, NY: Springer.

Sanfey, A. G. (2007). Social decision-making: Insights from game theory and neuroscience.

Science, *318*(5850), 598–602. http://doi.org/10.1126/science.1142996

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*(5626), 1755–1758. http://doi.org/10.1126/science.1082976

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. http://doi.org/10.1016/S1053-8119(03)00230-1

Schilbach, L., Bzdok, D., Timmermans, B., Fox, P. T., Laird, A. R., Vogeley, K., & Eickhoff, S. B. (2012). Introspective minds: using ALE meta-analyses to study commonalities in the neural correlates of emotional processing, social & unconstrained cognition. *PloS One*, *7*(2), e30920. http://doi.org/10.1371/journal.pone.0030920

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PloS One*, *6*(10), e23223. http://doi.org/10.1371/journal.pone.0023223

Schmitt, M. J., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The Justice Sensitivity Inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research*, *23*, 211–238. http://doi.org/10.1007/s11211-010-0115-2

Schmitt, M. J., Gollwitzer, M., Maes, J., & Arbach, D. (2005). Justice sensitivity: Assessment and location in personality space. *European Journal of Psychological Assessment*, *21*(3), 202–211. http://doi.org/10.1027/1015-5759.21.3.202

Schmitt, M. J., & Mohiyeddini, C. (1996). Sensitivity to befallen injustice and reactions to a real-life disadvantage. *Social Justice Research*, *9*(3), 223–238. http://doi.org/10.1007/BF02197249

Schmitt, M. J., Neumann, R., & Montada, L. (1995). Dispositional sensitivity to befallen injustice. *Social Justice Research*, *8*(4), 385–407. http://doi.org/10.1007/BF02334713

Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *The Journal of Neuroscience*, *27*(47), 12860–12867. http://doi.org/10.1523/JNEUROSCI.2496-07.2007

Schultz, W., Carelli, R. M., & Wightman, R. M. (2015). Phasic dopamine signals: From subjective reward value to formal economic utility. *Current Opinion in Behavioral Sciences*, *5*, 147–154. http://doi.org/10.1016/j.cobeha.2015.09.006

Schupp, H. T., Markus, J., Weike, A. I., & Hamm, A. O. (2003). Emotional Facilitation of Sensory Processing in the Visual Cortex. *Psychological Science*, *14*(1), 7–13. http://doi.org/10.1111/1467-9280.01411

Schwarz, L. A., & Luo, L. (2015). Organization of the locus coeruleus-norepinephrine system. *Current Biology*, *25*(21), R1051–R1056. http://doi.org/10.1016/j.cub.2015.09.039

Sebastian, A., Jung, P., Neuhoff, J., Wibral, M., Fox, P. T., Lieb, K., … Mobascher, A. (2016). Dissociable attentional and inhibitory networks of dorsal and ventral areas of the right inferior frontal cortex: a combined task-specific and coordinate-based meta-analytic fMRI study. *Brain Structure and Function*, *221*(3), 1635–1651. http://doi.org/10.1007/s00429-015-0994-y

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., … Greicius,

M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, *27*(9), 2349–56. http://doi.org/10.1523/JNEUROSCI.5587-06.2007

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *The Journal of Neuroscience*, *27*(18), 4826–2831. http://doi.org/10.1523/JNEUROSCI.0400-07.2007

Shackman, A. J., Salomons, T. V, Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, *12*(3), 154–67. http://doi.org/10.1038/nrn2994

Shamay-Tsoory, S. G. (2009). Empathic processing: Its cognitive and affective dimensions and neuroanatomical basis. In J. Decety & W. Ickes (Eds.), *The Social Neuroscience of Empathy* (pp. 215–232). Cambridge, MA: MIT Press.

Sheikh, S., & Janoff-Bulman, R. (2010). The "shoulds" and "should nots" of moral emotions: a self-regulatory perspective on shame and guilt. *Personality and Social Psychology Bulletin*, *36*(2), 213–224. http://doi.org/10.1177/0146167209356788

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286–1291.

Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*(4), 667–677. http://doi.org/10.1016/j.neuron.2010.07.020

Shweder, R. A., & Haidt, J. (1993). The future of moral psychology: Truth, intuition, and the pluralist way. *Psychological Science*, *4*(6), 360–365.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The “big three” of morality (autonomy, community, divinity) and the “big three” explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and Health* (pp. 119–169). New York, NY: Routledge.

Simon, J. J., Walther, S., Fiebach, C. J., Friederich, H.-C., Stippich, C., Weisbrod, M., & Kaiser, S. (2010). Neural reward processing is modulated by approach- and avoidance-related personality traits. *NeuroImage*, *49*(2), 1868–74. http://doi.org/10.1016/j.neuroimage.2009.09.016

Smetana, J. G. (2013). Moral development: The Social Domain Theory view. In P. D. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology, Vol 1: Body and Mind* (pp. 832–866). Oxford: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199958450.013.0029

Smith, E., Weinberg, A., Moran, T., & Hajcak, G. (2013). Electrocortical responses to NIMSTIM facial expressions of emotion. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *88*(1), 17–25. http://doi.org/10.1016/j.ijpsycho.2012.12.004

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced mutlilevel modeling* (2nd ed.). London, UK: SAGE Publications.

Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*(82), 1–7. http://doi.org/10.1038/s41562-017-0082

Sterling, P., & Eyer, J. (1988). Allostasis: A new paradigm to explain arousal pathology. In S. Fisher & J. Reason (Eds.), *Handbook of Life Stress, Cognition and Health* (pp. 629–649). Oxford, UK: John Wiley & Sons.

Sun, D., Chan, C. C. H., Hu, Y., Wang, Z., & Lee, T. M. C. (2015). Neural correlates of outcome processing post dishonest choice: An fMRI and ERP study. *Neuropsychologia*, *68*, 148–157. http://doi.org/10.1016/j.neuropsychologia.2015.01.013

Sussenbach, P., & Gollwitzer, M. (2015). Us(ed): The role of victim sensitivity in potentially exploitative intergroup relationships. *Group Processes & Intergroup Relations*, *18*(2), 241–255. http://doi.org/10.1177/1368430214556700

Süssenbach, P., & Gollwitzer, M. (2014). Us(ed): The role of victim sensitivity in potentially exploitative intergroup relationships. *Group Processes & Intergroup Relations*, *18*(2), 241–255. http://doi.org/10.1177/1368430214556700

Taber-Thomas, B. C., Asp, E. W., Koenigs, M., Sutterer, M., Anderson, S. W., & Tranel, D. (2014). Arrested development: Early prefrontal lesions impair the maturation of moral judgement. *Brain*, *137*, 1254–61. http://doi.org/10.1093/brain/awt377

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness. *Psychological Science*, *19*(4), 339–347. http://doi.org/10.1111/j.1467-9280.2008.02091.x

Talbot, C. F., Price, S. A., & Brosnan, S. F. (2016). Inequality responses in nonhuman animals. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of Social Justice Theory and Research* (pp. 387–403). New York, NY: Springer. http://doi.org/10.1007/978-1-4939-3216-0

Thayer, J., & Lane, R. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, *33*, 81–88.

Thierry, B. (2000). Covariation of conflict management patterns across macaque species. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 106–128). Berkeley, CA:

University of California Press.

Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, *94*(6), 1395–1415. http://doi.org/10.1119/1.1976413

Treadway, M. T., Buckholtz, J. W., Martin, J. W., Jan, K., Asplund, C. L., Ginther, M. R., … Marois, R. (2014). Corticolimbic gating of emotion-driven punishment. *Nature Neuroscience*, *17*(9), 1270–1275. http://doi.org/10.1038/nn.3781

Tyler, T. R. (1994). Psychological models of the justice motive: Antecedents of distributive and procedural justice. *Journal of Personality and Social Psychology*, *67*(5), 850–863. http://doi.org/10.1037/0022-3514.67.5.850

Van Bavel, J. J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology*, *6*, 167–172. http://doi.org/10.1016/j.copsyc.2015.08.009

Vermunt, R. (2014). *The Good, the Bad, and the Just*. Burlington, VA: Ashgate Publishing Company.

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *The Neuroscientist*, *20*(2), 150–159. http://doi.org/10.1177/1073858413494269

Wang, G., Li, J., Li, Z., Wei, M., & Li, S. (2016). Medial frontal negativity reflects advantageous inequality aversion of proposers in the ultimatum game: An ERP study. *Brain Research*, *1639*, 38–46. http://doi.org/10.1016/j.brainres.2016.02.040

Wassum, K. M., & Izquierdo, A. (2015). The basolateral amygdala in reward learning and addiction. *Neuroscience & Biobehavioral Reviews*, *57*, 271–283. http://doi.org/10.1016/j.neubiorev.2015.08.017

Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: the time-course of neural activity elicited by specific picture content. *Emotion*, *10*(6), 767–82. http://doi.org/10.1037/a0020242

Whalen, P. J. (2007). The uncertainty of it all. *Trends in Cognitive Sciences*, *11*(12), 499–500. http://doi.org/10.1016/j.tics.2007.08.016

Wijn, R., & Bos, K. van den. (2010). Toward a better understanding of the justice judgment process: the influence of fair and unfair events on state justice sensitivity. *European Journal of Social Psychology*, *40*, 1294–1301. http://doi.org/10.1002/ejsp

Wilson, B. J. (2012). Contra private fairness. *American Journal of Economics and Sociology*, *71*(2), 407–435.

Wu, H., Hu, X., & Fu, G. (2009). Does willingness affect the N2-P3 effect of deceptive and honest responses? *Neuroscience Letters*, *467*(2), 63–66. http://doi.org/10.1016/j.neulet.2009.10.002

Wu, Y., Leliveld, M. C., & Zhou, X. (2011). Social distance modulates recipient's fairness consideration in the dictator game: an ERP study. *Biological Psychology*, *88*(2–3), 253–62. http://doi.org/10.1016/j.biopsycho.2011.08.009

Yoder, K. J., & Decety, J. (2014a). Spatiotemporal neural dynamics of moral judgment: A high-density ERP study. *Neuropsychologia*, *60*, 39–45. http://doi.org/10.1016/j.neuropsychologia.2014.05.022

Yoder, K. J., & Decety, J. (2014b). The good, the bad, and the just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *Journal of Neuroscience*, *34*(12), 4161–4166. http://doi.org/10.1523/JNEUROSCI.4648-13.2014

Yoder, K. J., Harenski, C., Kiehl, K. A., & Decety, J. (2015). Neural networks underlying implicit and explicit moral evaluations in psychopathy. *Translational Psychiatry*, *5*(8), e625. http://doi.org/10.1038/tp.2015.117

Yoder, K. J., Lahey, B. B., & Decety, J. (2016). Callous traits in children with and without conduct problems predict reduced connectivity when viewing harm to others. *Scientific Reports*, *6*(February), 20216. http://doi.org/10.1038/srep20216

Young, L., Bechara, A., Tranel, D., Damasio, H., Hauser, M., & Damasio, A. (2010). Damage to ventromedial prefrontal cortex impairs judgment of harmful intent. *Neuron*, *65*(6), 845–51. http://doi.org/10.1016/j.neuron.2010.03.003

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(20), 8235–40. http://doi.org/10.1073/pnas.0701408104

Young, L., & Saxe, R. (2009). An FMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–405. http://doi.org/10.1162/jocn.2009.21137

Zanolie, K., De Cremer, D., Güroğlu, B., & Crone, E. A. (2015). Rejection in bargaining situations: An event-related potential study in adolescents and adults. *PLoS ONE, 10*(10), 1–13. http://doi.org/10.1371/journal.pone.0139953

Appendix E: Justice Sensitivity Inventory – Short Form

People react quite differently in unfair situations. How about you? For each statement, indicate how well it describes you.

1. It makes me angry when others are undeservingly better off than me.
2. It worries me when I have to work hard for things that come easily to others.
3. I am upset when someone is undeservingly worse off than others.
4. It worries me when someone has to work hard for things that come easily to others.
5. I feel guilty when I am better off than others for no reason.
6. It bothers me when things come easily to me that others have to work hard for.
7. I feel guilty when I enrich myself at the cost of others.
8. It bothers me when I use tricks to achieve something while others have to struggle for it.

Appendix F: Sensitivity to Punishment and Sensitivity to Reward Questionnaire

Please answer the following questions by clicking either yes or no.

1. Does the good prospect of obtaining money motivate you strongly to do some things?
2. Are you often afraid of new or unexpected situations?
3. Is it difficult for you to telephone someone you do not know?
4. Do you often do things to be praised?
5. Do you like being the center of attention at a party or a social meeting?
6. In tasks that you are not prepared for, do you attach great importance to the possibility of failure?
7. Are you easily discouraged in difficult situations?
8. Are you a shy person?
9. When you are in a group, do you try to make your opinions the most intelligent or the funniest?
10. Whenever possible, do you avoid demonstrating your skills for fear of being embarrassed?
11. Do you often take the opportunity to pick up people you find attractive?
12. When you are with a group, do you have difficulties selecting a good topic to talk about?
13. Do you generally give preference to those activities that imply an immediate gain?
14. Whenever you can, do you avoid going to unknown places?
15. Do you like to compete and do everything you can to win?
16. Are you often worried by things that you said or did?

17. Do you, on a regular basis, think that you could do more things if it was not for your insecurity or fear?
18. Do you sometimes do things for quick gains?
19. Comparing yourself to people you know, are you afraid of many things?
20. Do you often find yourself worrying about things to the extent that performance in intellectual abilities is impaired?
21. Do you often refrain from doing something you like in order not to be rejected or disapproved of by others?
22. Would you like to be a socially powerful person?
23. Do you often refrain from doing something because of your fear of being embarrassed?
24. Do you like displaying your physical abilities even though this may involve danger?

Appendix G: Justice Sensitivity Inventory

[Victim Sensitivity subscale, solitary subscale comprising JS-Self]

People react quite differently in unfair situations. How about you? First, we will look at situations to the advantage of others and to your own disadvantage[1].

1. It bothers me when others receive something that ought to be mine
2. It makes me angry when others receive a reward that I have earned
3. I cannot easily bear it when others profit unilaterally from me
4. It takes me a long time to forget when I have to fix others' carelessness
5. It gets me down when I get fewer opportunities than others to develop my skills
6. It makes me angry when others are undeservingly better off than me
7. It worries me when I have to work hard for things that come easily to others
8. I ruminate for a long time when other people are treated better than me
9. It burdens me to be criticized for things that are overlooked with others
10. It makes me angry when I am treated worse than others

[1] Participants indicated their response to each question using a 6-point scale (0 = "Not at all", 6 = "Exactly")

[Observer Sensitivity subscale, one of the three subscales comprising JS-Other]

Now, we will look at situations in which you notice or learn that someone else is being treated unfairly, put at a disadvantage, or used.

11. It bothers me when someone gets something they don't deserve
12. I am upset when someone does not get a reward he/she has earned
13. I cannot easily bear it when someone unilaterally profits from others
14. It takes me a long time to forget when someone else has to fix others' carelessness
15. It disturbs me when someone receives fewer opportunities to develop his/her skills than others
16. I am upset when someone is undeservingly worse off than others
17. It worries me when someone has to work hard for things that come easily to others
18. I ruminate for a long time when someone is treated nicer than others for no reason
19. It gets me down to see someone criticized for things that are overlooked with others
20. I am upset when someone is treated worse than others

[Beneficiary Sensitivity subscale, one of the three subscales comprising JS-Other]

Now, we will look at situations that turn out to your advantage and to the disadvantage of others.

21. It disturbs me when I receive what others ought to have
22. I have a bad conscience when I receive a reward that someone else has earned
23. I cannot easily bear it to unilaterally profit from others
24. It takes me a long time to forget when others have to fix my carelessness
25. It disturbs me when I receive more opportunities than others to develop my skills
26. I feel guilty when I am better off than others for no reason
27. It bothers me when things come easily to me that others have to work hard for
28. I ruminate for a long time about being treated nicer than others for no reason
29. It bothers me when someone tolerates things with me that other people are being criticized for
30. I feel guilty when I receive better treatment than others

[Perpetrator Sensitivity subscale, one of the three subscales comprising JS-Other]

Finally, we will look at situations in which you treat someone else unfairly, discriminate against someone, or exploit someone.

31. It gets me down when I take something from someone else that I don't deserve
32. I have a bad conscience when I deny someone the acknowledgment he or she deserves
33. I cannot stand the feeling of exploiting someone
34. It takes me a long time to forget when I allow myself to be careless at the expense of someone else
35. It disturbs me when I take away from someone else the possibility of developing his or her potential
36. I feel guilty when I enrich myself at the cost of others
37. It bothers me when I use tricks to achieve something while others have to struggle for it
38. I ruminate for a long time when I treat someone less friendly than others without a reason
39. I have a bad conscience when I criticize someone for things I tolerate in others
40. I feel guilty when I treat someone worse than others

Appendix H: Interpersonal Reactivity Index – Modified

The following statements inquire about your thoughts and feelings in a variety of situations. Indicate how well each item describes you by choosing the appropriate number on the scale at the top of the page: 1, 2, 3, 4, or 5. When you have decided on your answer, fill in the number in the blank next to the item. *Read each item carefully before responding.* Answer as honestly and as accurately as you can.

Thank you.

1. I often have tender, concerned feelings for other people less fortunate than me.
2. I sometimes find it difficult to see things from the "other guy's" point of view.
3. Sometimes I don't feel very sorry for other people when they are having problems.
4. In emergency situations, I feel apprehensive and ill-at-ease.
5. I try to look at everybody's side of a disagreement before I make a decision.
6. When I see someone being taken advantage of, I feel kind of protective towards them.
7. I sometimes feel helpless when I am in the middle of a very emotional situation.
8. I sometimes try to understand my friends better by imagining how things look from their perspective.
9. When I see someone get hurt, I tend to remain calm.
10. Other people's misfortunes do not usually disturb me a great deal.
11. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments.
12. Being in a tense emotional situation scares me.
13. When I see someone being treated unfairly, I sometimes don't feel very much pity for them.
14. I am usually pretty effective in dealing with emergencies.
15. I am often quite touched by things that I see happen.

16. I believe that there are two sides to every question and try to look at them both.

17. I would describe myself as a pretty soft-hearted person.

18. I tend to lose control during emergencies.

19. When I'm upset at someone, I usually try to "put myself in his shoes" for awhile.

20. When I see someone who badly needs help in an emergency, I go to pieces.

21. Before criticizing somebody, I try to imagine how I would feel if I were in their place.

Appendix I: Instructions and questions used in the Knowledge Misrepresentation Task

Below are sample questions from 8th grade textbook test banks on the topics of history, geography and algebra. The questions are rather simplistic, but even so they provide an option for honest but uneducated students to mark "I don't know".

We have hypotheses about the effects of having an "I don't know" option on these questions. It is difficult to test the hypotheses with educated adults, but **in the event that you are unable to solve or answer any of the following questions, please feel free to indicate so by marking "I don't know".**

1. [***Unanswerable***] Who signed the Détente of Bruges?
    a. James Monroe
    b. William King
    c. Benjamin Franklin
    d. Franklin D. Roosevelt
    e. I don't know

2. [***Unanswerable***] (3x - 5)Þ = 13
    a. x = 2.6
    b. x = 28
    c. x = 13
    d. x = 5
    e. I don't know

3. Florida is an example of
    a. an isthmus
    b. a plateau
    c. a peninsula
    d. an island
    e. I don’t know

4. What religion is practiced by most people who live in India?
    a. Confucianism
    b. Hinduism
    c. Christianity
    d. Buddhism
    e. I don’t know

5. Under which of the following circumstances would you be most likely to find snow in equatorial regions?
    a. in areas below sea level
    b. in areas at high latitudes
    c. in areas at high elevations
    d. in winter
    e. I don’t know

6. If *m* represents the total number of months that Jill worked, and *p* represents Jill's average monthly pay, which of the following expressions represents Jill's total pay for the months she worked?

    a. $m + p$
    b. $m \div p$
    c. $m\ x\ p$
    d. $m - p$
    e. I don't know

7. A car can seat *c* adults. A van can seat 4 more than twice as many adults as the car can. In terms if *c*, how many adults can the van seat?

    a. c + 8
    b. c + 12
    c. 2c – 4
    d. 2c + 4
    e. I don't know

8. Which of the following equations has the same solution as the equation 2 x + 6 = 32?

    a. 2 (x + 3) = 32
    b. 2x = 38
    c. 2 (x – 3) = 16
    d. x – 3 = 16
    e. I don't know

9. The purpose of the Underground Railroad was to
    a. make transportation to the western frontier easier
    b. provide an escape route for slaves fleeing their masters
    c. take Native Americans to reservations in the West
    d. provide a way for European immigrants to move to the Midwest
    e. I don't know

10. The first permanent English settlement in North America was
    a. St. Augustine
    b. Santa Fe
    c. New Amsterdam
    d. Jamestown
    e. I don't know

11. Near the end of the Second World War, the United States military dropped atomic bombs on cities in
    a. Japan
    b. China
    c. Germany
    d. the Soviet Union
    e. I don't know

12. The United States fought one of its longest wars, from 1964 to 1973, in

a. Vietnam

b. Germany

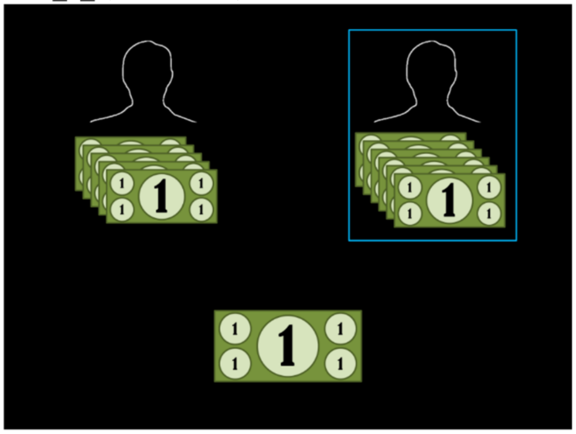
c. Kuwait

d. Japan

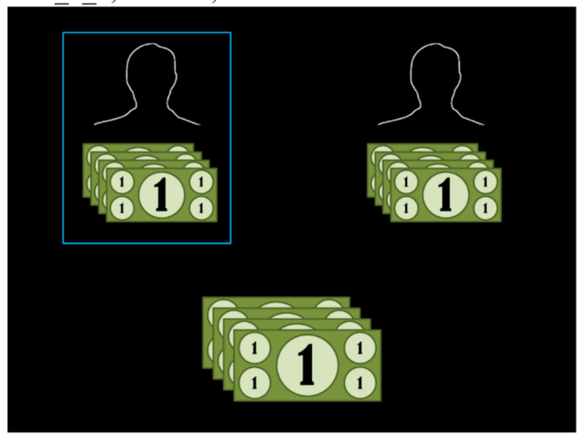
e. I don’t know

Appendix J: Example stimuli from Three-party Distribution Task used in Chapters 4 and 5

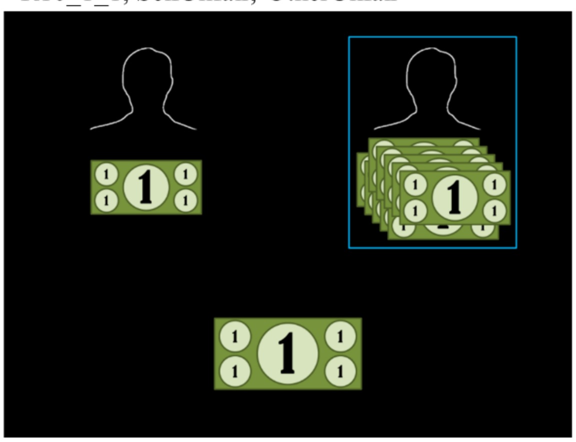
R6_5_1, SelfUnfair, OtherFair



L4_4_4, SelfFair, OtherFair



R10_1_1, SelfUnfair, OtherUnfair



L6_1_5, SelfFair, OtherUnfair