Title
Social Decision-Making and Close Others

Author
Guassi Moreira, Joao F

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Social Decision-Making and Close Others

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Psychology

by

João F. Guassi Moreira

2022

ABSTRACT OF THE DISSERTATION

Social Decision-Making and Close Others

By

João F. Guassi Moreira

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2022

Professor Jennifer A. Silvers, Chair

Human beings are inextricably linked to others. The ties we share with other people throughout the lifespan are a fundamental aspect of our existence. It follows that any decision we make as individuals has the potential to affect others, including those closest to us. This notion that our decisions frequently have consequences for others has fascinated psychologists for decades. Recently there has been a resurgence of interest in the psychological community on the subject of social decision-making, defined as decisions that have some kind of implicit or explicit social consequence. This surge has carried with it an unprecedented understanding of social decision-making behaviors, yet social decision-making studies are frequently constrained by a lack of ecological validity, specifically in terms of social decision partners (e.g., pairing participants with strangers, confederates, or fictive individuals). Doing so means that much of the literature may not be applicable to the most frequent and impactful decisions humans make: those involving close others. I aimed to help resolve these limitations in this dissertation. Across three studies I examined social decision preferences between parents and friends, two developmentally important relationships, at the behavioral, neural, and cognitive levels. Study 1 shows that, overall, adolescents appear to exhibit a distinct preference towards parents over friends when making social decisions involving conflicting outcomes for these two close others. Study 2 expands this work by showing that value-based neural representations of these close others is associated with behavioral preference; Study 3 continues this conceptual line of work by broadening the scope to show that value-based cognitive representations—derived from naturalistic text data—are also linked to behavioral social decision preferences. The contributions of this work towards broader theories of social decision-making, developmental science, and social neuroscience are discussed.

The dissertation of João F. Guassi Moreira is approved.

Adriana Galván

Andrew J. Fuligni

Carolyn Parkinson

Tor D. Wager

Jennifer A. Silvers, Committee Chair

University of California, Los Angeles

2022

This dissertation is dedicated the Guassi and Moreira families, whose sacrifices across generations have given me the opportunities to pursue my passion.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

FIGURES

Acknowledgments

Participating in a doctoral program and completing a dissertation takes a while. It is even longer when you consider the supporting groundwork that has to be laid before a student even attends college (let alone graduate school). I feel quite fortunate to have been supported by several generous close others who care fiercely about my wellbeing. Foremost, I am *deeply* indebted to my parents, Marco and Sandra, who have made countless personal sacrifices so I could follow my passions from a young age. Without you two, none of this is possible. My sister, Ana, has been a bright light in my life—your unconditional love and positivity have helped keep me together since the day you were born. Relatedly, I feel humbled when reflecting on how strongly my family in Brasil has supported me from afar—I have always felt the warmth of your love and support.

In addition to my biological family, I am fortunate to have an incredible extended family comprised of several close friends that I made from childhood through college. Naming and thanking you all would exceed the length of this dissertation, so I can only focus on a few. I most certainly would not have accomplished this feat without my best friend, Kevin McCannon, who has been a constant source of support for me dating back to our caddie shack days all the way through my graduate school sojourns to desert. I also would not be here without the help of Max Savinkov, though I mean this here in quite a literal sense since he was the one who helped move me to Los Angeles. Among my friends from my undergraduate days, Julia Monk and Vyshnav Kakivaya stand out as the ones that most encouraged my career and scientific pursuits, albeit in their own unique ways—thank you for sticking around. I owe a debt of gratitude to Michelle Nussey, who at times carried me when I did not have the strength to carry myself. Finally, I thank Westley Banks. I miss being able to call you for your cool, calm, and thoughtful reflections whenever I needed them. I wish you could be here to see this.

I am doubly fortunate to say that this ‘extended family’ has grown tremendously during graduate school, where I have made several lifelong friends. Michael Perino, my academic day one, thank you for your mentorship and, more importantly, your friendship. I am not sure I deserved a graduate mentor as invested and thoughtful as you, but I am sure as hell grateful you came along (despite your odd penchant for blueberry beer). Sarah Tashjian, words cannot fully begin to express my gratitude for you. You have been the perfect colleague, collaborator, mentor, and friend for me at this stage of my career and life. I have learned so much from you—you are truly like a big sister to me. I would share some of my favorite memories but I am afraid I already do that too often when I spend time with you and Joel (though perhaps I can quickly paraphrase one of my favorites: ‘something something swap the labels on this conversation, something something couldn’t tell who is who’). Emilia Ninova, your calm demeanor and grounded advice kept me sane during my first two years in graduate school (even if I sometimes had to listen while lying face down on the ground in A191); I look forward to ending up in New York with you and John, just as we planned it. Razia Sahi, I so cherish our affective science talks, horror movie nights, summer cookouts, and Halloween house crawl cigarette breaks; your positivity has helped give me confidence and perspective when I needed it most. Adriana Méndez Leal, my first SANDLab sibling, thank you for being there for me at times when it felt like I was at my absolute weakest; I could not have weathered some serious storms without your perspective. Natalie Saragosa-Harris, your chingona energy was always a constant source of inspiration for me. Wesley Meredith, your supportive and thoughtful vibes during our ‘homie chats’ were much appreciated during these finals chapters. Pooja Patel, I hope the future brings

us many more opportunities to drink wine, smoke cigars, play bar trivia, and quote our favorite Succession lines. Gwen Price, your constant stream of kitty content made sure my dopamine levels never crashed entirely (also, no toucha!) Elisa Baek, you checked into the game a little later, but you could not have come at a better time. I look forward to seeing you a lot more in this next chapter.

In addition to friends and family, I could not have completed this dissertation without a fiercely supportive group of mentors. I am so grateful that Dr. Eva Telzer gave me my first home in science and guided me through my undergraduate years. From there, I was incredibly fortunate to have received the support of such a talented, brilliant, and caring group of secondary faculty mentors: Drs. Adriana Galván, Amanda Montoya, Andrew Fuligni, Bridget Callaghan, Carolyn Parkinson, and Matt Lieberman—I have felt that each of you cared about my growth as a scientist and looked out for me in your own unique ways. I am all the better for it.

My advisor, Dr. Jennifer Silvers, once remarked early in my doctoral career that she felt she had 'hit the academic jackpot' for having been able to found her lab with Emilia and myself in the fold. Six years later, I have to say I am the one who hit the jackpot. Jen, from day one, I have always felt that you were an advocate for me, and that you cared not only about my career, but also about my wellbeing as a human being. I could not have asked for a better doctoral advisor, mentor, and friend—all in one—for this stage of my life. I deeply appreciate your patience in guiding me through this process and your many pushes for me to think deeper, write clearer, and communicate more effectively. More importantly, I am even more grateful for the life advice you imparted unto me during this process—your mentorship helped me grow into myself as a person. I will undoubtedly look back on the time spent working with you as one of the most fulfilling and gratifying eras of my life. It has been an honor.

Finally, I have John Flournoy and Zachary Horne to thank for helping me navigate Bayesland while analyzing these data.

Research in this dissertation was funded by a National Science Foundation Graduate Research Fellowship, a National Science Foundation CAREER award (J.A. Silvers), a National Institutes of Health Predoctoral T32 fellowship, generous funds from the University of California, Los Angeles, and the American Psychological Foundation.

Select Vita

EDUCATION

May 2020	**University of California, Los Angeles** Advanced to Doctoral Candidacy, Developmental Psychology
2016-2017	**University of California, Los Angeles** Master of Arts, Developmental Psychology
2013 – 2016	**University of Illinois, Urbana-Champaign** Bachelor of Science, Psychology Magna Cum Laude with Distinction

HONORS, AWARDS, AND FUNDING

2021-2022	Dissertation Year Fellowship - UCLA
2019	SfN Graduate Travel Student Award - BRI-Semel Institute (UCLA)
2019	Travel Award, NIMH ABCD Workshop
2019-2020	Brain & Behavior Development during Adolescence T32 Fellowship - NICHD (UCLA)
2018	APA Early Graduate Student Researcher Award
2018	Pre-Registration Challenge, Center for Open Science
2018	Travel Award, Intl. Society for Developmental Psychobiology
2018	Travel Award, Flux Congress
2018	Poster Spotlight, Society for Affective Science
2017-2018	Dr. Ursula Mandel Fellowship - UCLA
2017	Graduate Summer Research Mentorship –UCLA
2017	Poster Award, Social Affective Neuroscience Society
2017	Multimodal Neuroimaging Training Program (MNTP) Travel Award - U. Pittsburgh/Carnegie Mellon
2016-2021	NSF Graduate Research Fellowship -
2016-2017	Graduate Dean's Scholar Award - UCLA
2016	APAGS/Psi Chi Junior Scientist Fellowship -
2016	Competitive Edge Graduate Research Summer Program – UCLA
2016	Spring/Early Summer Conference Travel Grant – U. Illinois
2015	Summer Undergraduate Research Fellowship - U. Illinois
2014	Independent Funding (Experiment.com)
2014-2016	James Scholar (Honors Program) – U. Illinois
2014-2016	Dean's List – U. Illinois
2014-2015	Certificate of Excellence – Access & Achievement Program U. Illinois
2013-2016	President's Award Scholarship - U. Illinois
2013	Great Lakes Credit Union Educational Scholarship -

SELECTED PUBLICATIONS

Guassi Moreira, J.F., Méndez Leal, A.S., Waizman, Y.H., Saragosa-Harris, N.M., Ninova, E., & Silvers, J.A. (accepted). Early caregiving adversity differentially shapes behavioral sensitivity to reward and risk during decision-making. *Journal of Experimental Psychology: General.*

Guassi Moreira, J.F., Méndez Leal, A.S., Waizman, Y.H., Saragosa-Harris, N.M., Ninova, E., & Silvers, J.A. (2021). Revisiting the neural architecture of adolescent decision-making: Univariate and multivariate evidence for system-based models. *Journal of Neuroscience, 41*(28), 6006-6017. doi: 10.1523/JNEUROSCI.3182-20.2021

Guassi Moreira, J.F., McLaughlin, K.A., & Silvers, J.A. (2021). Characterizing the network architecture of emotion regulation neurodevelopment. *Cerebral Cortex*, *31*(9), 4140-4150. doi: 10.1093/cercor/bhab074

Waizman, Y.H., Méndez Leal, A.S., **Guassi Moreira, J.F.,** Saragosa-Harris, N.M., Ninova, E., & Silvers, J.A. (in press). Emotion regulation strategy usage explains links between institutional caregiving and elevated internalizing symptoms. *Developmental Psychobiology,* advance online publication. doi: 10.1002/dev.22078

Guassi Moreira, J.F., Tashjian, S.M., Galván, A.G., & Silvers, J.A. (2021). Computational and motivational mechanisms of human social decision making involving close others. *Journal of Experimental Social Psychology, 93*, 1-10. doi: 10.1016/j.jesp.2020.104086

Guassi Moreira, J.F., Sahi, R., Ninova, E., Parkinson, C., & Silvers, J.A. (2020). Performance and belief-based emotion regulation capacity and tendency: Mapping links with cognitive flexibility and perceived stress. *Emotion,* advance online publication. doi: 10.1037/emo0000768

Guassi Moreira, J.F., Tashjian, S.M., Galván, A., & Silvers, J.A. (2020). Is social decision making for close others consistent across domains and within individuals? *Journal of Experimental Psychology: General, 149*(8), 1509-1526. doi: 10.1037/xge0000719

Guassi Moreira, J.F., McLaughlin, K.A., & Silvers, J.A. (2019). Spatial and temporal cortical variability track with age and affective experience during emotion regulation in youth. *Developmental Psychology, 55*(9), 1921-1937. doi: 10.1037/dev0000687

Silvers, J.A., & **Guassi Moreira, J.F.** (2019). Capacity and tendency: A neuroscientific framework for the study of emotion regulation. *Neuroscience Letters, 693*, 35-39. doi: 10.1016/j.neulet.2017.09.017

Tashjian, S.T., **Guassi Moreira, J.F.,** & Galván, A. (2019). Amygdala pattern response similarity to faces relates to trust bias in adolescents. *Journal of Cognitive Neuroscience, 31*(11), 1726-1741. doi: 10.1162/jocn_a_01448

Guassi Moreira, J.F., Tashjian, S.M., Galván, A., & Silvers, J.A. (2018). Parents versus peers: Assessing the impact of social agents on decision making in young adulthood. *Psychological Science, 29*(9), 1526-1539. doi: 10.1177/0956797618778497

General Introduction

Human beings are inextricably linked to others. The ties we share with other people throughout the lifespan are a fundamental aspect of our existence. It follows that any decision we make has the potential to affect others, including those closest to us. The notion that our decisions frequently have consequences for others has fascinated psychologists for decades (Festinger, 1943; Milgram, 1963; Torrance, 1957). More recently, the psychological science community has seen a resurgence of interest on the subject of *social decision-making*, defined as decisions that have some kind of implicit or explicit social consequence. In this dissertation, I briefly highlight the importance of social decision-making, broadly survey a subset of the literature's most popular and impactful findings, identify two critical limitations of prevailing social decision-making research protocols, and then propose a series of studies that aim to help resolve these limitations. I then motivate each of the studies in detail and justify their proposed methodologies.

The Importance of Social Decision-making

On a daily basis, we make hundreds of decisions, many of them possessing social ramifications: Who should I spend time with on the weekend? What kind of birthday gift should I get for my spouse? Will I ask my colleague for help on a project? Should I loan my friend money? Such decisions have both macro- and micro-level consequences. At the macro level, as these decision scenarios become interlinked with each other—the outcome of previous decisions influencing subsequent decisions—they begin to layer and form the foundation of individuals' social trajectories and collectively, our society (Henrich, 2002; Mathews, 1987). At the micro level, social decision-making scenarios are replete with a rich diversity of psychological processes, encompassing phenomena such as trust, generosity, altruism, revenge, guilt, in addition to a myriad of others. It is not difficult to understand why psychologists have held a long-standing interest in social decision-making. Not only are social decisions interesting in their own right, but they provide a vehicle for understanding phenomena that have been at the center of centuries-old philosophical debates and are highly consequential for shaping everyday outcomes that involve how individuals spend their time, effort, and money.

Surveying the Social Decision-Making Literature

The past fifteen years have seen a surge in interest in the study of social decision-making. The byproduct of increasing interdisciplinarity between fields such as psychology, economics, and neuroscience, this recent wave of social decision-making literature has been marked by unprecedented sophistication. Quantitative insights from economics have allowed the field to begin to mathematically operationalize complex psychological processes (Camerer, 2011; Lewin, 1938), affording an avenue for unambiguous, formal theories. Neuroimaging techniques from cognitive neuroscience provide *in vivo* recording of brain activity, potentially resolving questions of mechanism that cannot be answered solely with traditional behavioral assessments. Coupled with psychology's classic methodological tenets of strong, tightly controlled experiments, a deep collection of knowledge on a range of social decision contexts has been amassed. Below, I survey two particularly well-studied types of social decision-making: prosocial behavior and resource allocation/reciprocal interactions. In focusing on these two specific forms of social decision-making, I aim not to summarize the entire field, but instead showcase two exemplary lines of social decision-making research that are highly applicable to everyday life and carry the flexibility to address many related but distinct topics of study in psychology.

Prosocial Behavior. Prosocial behavior is one of the most well-studied forms of social decision-making. Debate exists within numerous fields about how to define prosocial behavior and its relationship with related constructs such as altruism (Eisenberg et al., 1998; Kerr et al., 2004; Rhoads et al., 2021). For my purposes, I broadly define prosocial behavior as deliberate actions that benefit another individual or entity, typically accompanied with the intention of doing so. Scholarly attention towards prosocial behavior stems from questions about human nature posed millennia ago by philosophers (Annas, 1993; Strohminger & Nichols, 2014). Prosocial behavior, especially when it comes at a cost to the decision maker, was thought to reflect information about inherent moralistic tendencies. More recently, prosocial behavior has been a focus of inquiry because of its evolutionarily relevance (Kerr et al., 2004), associations with mental health (Aknin et al., 2013; Weinstein & Ryan, 2010; Armstrong-Carter, Guassi Moreira, Ivory, & Telzer, 2020), and role in promoting positive relationships (e.g., family dynamics in adolescence; Telzer, Masten, Berkman, Lieberman, & Fuligni, 2010). Many of the studies on these facets and correlates of prosocial behavior have taken a developmental approach because it is theorized the psychological substrates of prosocial behavior undergo sensitive periods during juvenile development (Do et al., 2017).

Much of what is known about prosocial behavior comes from research employing helping tasks, wherein participants have the opportunity to provide some form of assistance to another individual. Research in non-human animal models and young children has revealed that prosocial behaviors are both evolutionarily and ontogenetically conserved (Ben-Ami Bartal et al., 2011, 2014; Knafo et al., 2011; Vaish et al., 2016). In fact, several studies have boldly suggested that prosocial behavior is so deeply woven into the fabric of human social decision-making that individuals ‘default’ to prosocial decisions when decisions must be made quickly and intuitively (Carlson, Aknin, & Liotti, 2016; Zaki & Mitchell, 2013). Early neuroimaging work on this topic showed that brain systems typically implicated in social cognition and reward processing may undergird prosocial behavior (Do et al., 2017; Telzer et al., 2011; Zaki et al., 2012). These advances laid the groundwork for increasingly sophisticated approaches that paired computational reinforcement learning models with functional magnetic resonance imaging (fMRI) to strongly suggest that mechanisms of prosocial behavior lie in one's ability to use social cognitive skills to track value-based probabilistic contingencies that affect others (Lockwood et al., 2016). This is relevant because it implies that prosocial motivations co-opt core evaluation processes such as valuation, speaking to the importance of prosocial motivations among the hierarchy of psychological processes.

Critically for social decision-making research, it is apparent from decades of literature that prosocial tendencies vary both between and within individuals. Prosocial behaviors vary between-persons (Contreras-Huertas et al., 2020; Kwak et al., 2014; Lockwood et al., 2016) for at least two major, related, reasons. First, prosociality itself is trait-like and is related to core personality traits (e.g., Big Five)[1], and we know from a rich history of personality psychology that personality traits vary drastically among individuals. Second, and relatedly, prosociality is thought to have specific genetic underpinnings. To the extent the relevant genotypes vary, so will prosocial behavior (Conway & Slavich, 2017; Israel et al., 2015; Padilla-Walker & Fraser, 2014; Thielmann et al., 2020). Longitudinal work in the developmental literature suggests prosocial behavior changes within-individuals across the lifespan as a function of age, experience, and

[1] It is important to acknowledge prosociality itself is often conceptualized as a personality trait, either as a distinct trait or a pole of a more general trait (e.g., agreeableness) depending on the theoretical disposition (Caspi et al., 2005). For the purpose of this dissertation, I am uniquely concerned with prosocial decision behavior as a type of social decision-making, as opposed to the totality of prosociality as a broader construct (which contains other elements beside decision behavior).

other developmental factors (Do et al., 2017; Fu et al., 2017; Kwak & Huettel, 2016; Mayr & Freund, 2020; Padilla-Walker et al., 2013). Prosocial behavior also appears to be highly context-dependent, both between and within individuals. For example, individuals are less likely to engage in prosocial behavior if doing so requires increased effort (Lockwood et al., 2017), or if one's prosocial behavior is likely to benefit out-group members (Dunham, 2018; Padilla-Walker et al., 2013). Together, these findings begin to highlight that social decision behaviors can vary quite dramatically based on the decision-maker, their age, and other contextual features. Variability of this degree implies that there are both within- and between-person factors that can act as determinants of prosocial decisions. In turn, this suggests the manipulation of such variables can be fruitful for identifying mechanisms and boundary conditions involved in prosocial decision-making. What is specifically intriguing is that prosocial behaviors appear to vary as a function of the target involved (e.g., ingroup vs outgroup member, familiar versus unfamiliar other, etc.), suggesting social decision tendencies must be thought of as granular, rather than domain-general. More broadly, it also requires acknowledgement that experimental protocols must have flexible protocols that can account for the intrinsic variability of social decision-making.

Resource Allocation and Reciprocal Interactions. Another widely studied form of social decision-making involves the allocation of resources and reciprocal interaction[2]. Because an agent's future behavior is a function of past experiences and motivations (Lewin, 1938), humans must reason critically about how their actions may affect others' behavior in the future (e.g., "If I am generous with John now, will he be more likely to be generous with me at a future time?").

[2] Though these two types of social decisions are conceptually distinct, researchers frequently pair them together in experiments and thus I review them in tandem.

Such decisions depend markedly on a decision maker's tendencies towards trust, generosity, harm, fairness, and inequity (Crockett et al., 2014; Gonzalez & Chang, 2019; Hackel & Zaki, 2018; Rilling & Sanfey, 2011). Decisions about resource allocation and reciprocity form the basis of positive interpersonal dynamics, and are associated with longitudinal increases in wellbeing (Poulin & Haase, 2015; J. A. Simpson, 2007). Looking beyond small-scale interactions, cultural and societal norms involving trust and reciprocity can have sweeping consequences (Avdeenko & Gilligan, 2015; Johnson & Mislin, 2011; Kimbrough & Vostroknutov, 2016; Schnakenberg & Turner, 2019, 2021). For example, aggregation of individual social decision tendencies within a nation is thought to explain why countries whose citizens exhibit high levels of trust show greater per-capita GDP, higher quality government bureaucracies, and more efficient judicial systems (Johnson & Mislin, 2011). Even individual interactions can have broader, compounding social effects, such as politicians deciding whether to obtain useful information from lobbyists at the risk of appearing corrupt (Schnakenberg & Turner, 2019). Not only does this collectively signal the importance of studying the role that trust and reciprocity play in social decision-making, but it also highlights the significance of such work for informing multiple arcs of research within psychology and beyond.

Experiments involving resource allocation and reciprocal interactions typically use turn-based tasks that involve some kind of alternating resource distribution between participants (Colin F. Camerer, 2015; Johnson & Mislin, 2011). More complex designs may involve multi-phase experiments where individuals learn about particular social agents and are then allowed to interact with them in some way (Hackel & Zaki, 2018). Several decades of psychological studies show that humans have a strong preference to minimize harm unto others during resource allocation, even going so far as to disobey perceived and objective authority figures (Milgram, 1963). This property of human decision-making appears to underlie individuals' aversion to inequity when allocating resources between themselves and others, potentially fueling cooperative behavior (Pärnamets et al., 2019; Yoder & Decety, 2018). In fact, this goal is rooted so deeply that rewards for oneself made from harming others are represented in the brain in a way that is fundamentally different from other types of rewards—these 'tainted' rewards result in so-called 'corruptions' of neural value signals (Crockett et al., 2017)—and people engage in motivated misremembering in order preserve positive self-images (Carlson et al., 2020).

Beyond wishing to minimize harm to others, humans also appear to possess a strong desire to maximize reciprocity. Classic studies using economic tasks such as the public goods game or the ultimatum game have shown individuals readily share resources with others who share resources with them (Feldmanhall & Chang, 2018; Grecucci et al., 2013; Rilling & Sanfey, 2011; van Hoorn et al., 2016). In fact, reciprocity behaviors are so salient that neuroimaging and computational modeling studies suggest that observations of others' reciprocity behavior are given priority when encoding social representations at the neural level and are even weighted as heavily as value-based predictions (e.g., expected value of a given decision alternative) (Hackel, Doll, & Amodio, 2015; Hackel & Zaki, 2018). Collectively these studies highlight how social decision-making in contexts involving reciprocity and resource allocation underlies everyday human interaction, as well as higher-level, societal dynamics (see Johnson & Mislin, 2011).

Just as with prosocial behavior, this literature is also marked by considerable heterogeneity in decision tendencies. While most people are averse to harming others, some individuals are more willing than others to behave selfishly during resource allocation. In fact, a subset of individuals even show an *overall* propensity to harm and take advantage of others in these contexts for self-advancement (Crockett et al., 2017; Paulhus, 2014). Similarly, another subset of individuals are less willing to outright harm others but are nevertheless highly inclined to behave selfishly and buck collective rewards (Rilling & Sanfey, 2011). These findings have contributed to the emergence of a dual systems perspective suggesting that selfishness and cooperation—both at the between and within-person levels—duel each other in a push-pull dance to sway cooperation and reciprocation tendencies (Pärnamets et al., 2019). While this possibility entices ancient and modern intuitions of the human psyche (Pfeifer & Allen, 2012), it is likely that other moderating variables exist to explain heterogeneity in social decision-making tendencies in these contexts. Such heterogeneity implies there are rich avenues to further add depth to the literature, both in terms of individual differences but also in contextual specificity of main effects. In the next section I introduce a set of features common to most social decision-making that currently place a bottleneck on the boundaries of our knowledge and then solutions that the proposed dissertation will seek to contribute.

Bottlenecks to the Utility of Current Social Decision-Making Research

Despite making tremendous strides, the literature on social decision-making is far from complete. I argue that social decision-making research to date has been necessarily 'deep' at the expense of being 'wide'. That is, extant research has focused its efforts on thoroughly examining phenomena of interest under a narrow set of conditions in the interest of bolstering internal validity, or the degree to which one can claim an experimental effect is actually driven by its purported cause. This strategy has afforded the field experimental power to develop some of the most quantitatively articulate and unambiguous theories in all of psychological science (C.F. Camerer, 2011), leading to clear and elegant mathematic definitions of complex psychological phenomena (such as guilt, for instance (Chang et al., 2011)). This benefit cannot be overstated: the history of psychology has been rife with 'soft', verbal theories that eventually become scientifically untenable and misguiding to research (Meehl, 1978). As such, existing social decision-making research deserves commendation for advancing a line of work that is quantitatively precise and scientifically tenable. However, focusing on internal validity often comes with the trade-off of sacrificing external validity (investigating 'widely') and therein arise the gaps the current work seeks to address.

In assaying the social decision-making literature, it appears much of the literature has poor external validity. External validity, which refers to the ability of a study's results to remain applicable to psychological phenomena outside of the laboratory and generalize to novel contexts, is a critical component of psychological science (Tebes, 2000). If our experiments cannot apply to everyday life and meaningfully capture naturally encountered psychological conditions, then they are virtually useless at best and unscientific at worst. In fact, scholars speculate poor external validity to be one of the root causes of psychology's recent replication crisis, further underscoring the notion that scientific work must be externally valid if it is to be rigorous (Open Science Collaboration, 2015; McCarthy et al., 2018; Simons, Shoda, & Lindsay, 2017). Here I identify two specific pitfalls in the social decision-making literature that limit the field and have motivated the proposed work: (i) lack of work involving social agents who are close others and (ii) lack of experimental scenarios that consider the impact of one's decisions on multiple others, especially in conflicting ways. The following sections review each of these pitfalls in detail, argue why they are important issues to address, and enumerate the benefits of addressing each.

Close Others as Social Agents. Historically, the vast majority of social decision-making research has relied on experimental protocols that pair participants with fellow participants, or a confederate whom they are led believe are also naïve participants. Reasons for these practices are twofold. First, most prior social decision-making research has been focused on decisions involving anonymous or unfamiliar social agents (Crockett et al., 2014; Kappes et al., 2018). Second, other research questions have been nominally agnostic to the relationship between the participant and the social agents affected by their decision, but want to ensure that closeness with a social agent does not act as a confounding variable (C.F. Camerer, 2011; Colin F. Camerer, 2015; Krajbich et al., 2009).

Studying social decision-making regarding unfamiliar others is not inherently weak, but I argue that *purely* using designs with anonymous strangers or confederates is limiting. Close relationships not only influence behaviors in ways that other relationships simply cannot, but outcomes involving these others also tend to be more consequential in our everyday lives (Aron et al., 1991; Meyer & Anderson, 2000; Telzer et al., 2010). Consider the literature on prosocial behavior as an illustrative example—the majority of studies in this literature have examined prosociality as it relates to distant, unfamiliar others. Yet, evidence suggests that minimal group membership can change prosocial tendencies (Do & Telzer, 2019), and rewards for close others are valued much more highly distant others (e.g., unfamiliar individuals, acquaintances) (Braams & Crone, 2016; Feng et al., 2013; S.-M. Wang et al., 2019; Zhao et al., 2014). Failure to help, or even choosing to harm, an unfamiliar other for personal gain might help answer long-debated philosophical questions about human nature, but likely has little bearing on one's day-to-day life. Choosing to harm a close other for personal gain (e.g., cheating on a spouse), on the other hand, could have far-reaching consequences. Thus, focusing on confederates and anonymous partners has likely diminished the ecological relevance of much recent work on social decision-making, despite its use of rigorous methods and high internal validity.

Including close others as agents in social decision-making experiments will boost the external validity of the literature and help scientists make predictions about scenarios that frequently color human life (Berkman & Wilson, 2020). However, there also other benefits to the field of psychology as a whole. Social decision preferences involving a close other can lend general insight into the nature of those close relationships, such as indicating whether a given close relationship is more privileged than another and under what circumstances. For instance, perhaps individuals prioritize parents over other close relationships involving social decisions with monetary outcomes as a form of reciprocity, while prioritizing friends over other close relationships when social outcomes are at stake due to better fulfillment of socioemotional needs. Differences between individuals' social preferences involving close others during decision-making may also help corroborate theories of personality, or even serve as a test of whether certain decision-related cognitions are truly domain general or varying by social context.

Navigating Trade-Offs Involving Multiple Individuals. Another limitation of the extant social decision-making work is that most experiments involve making decisions that only affect one social agent. Slight variations of this practice exist – participants may sometimes engage with multiple partners over the course of a session despite making decisions that only affect one person at a time (Huck et al., 2012), they may choose between multiple partners with whom they want to complete a task with (Delgado et al., 2005; Hackel & Zaki, 2018), or they may interact with large groups (Hackel et al., 2017). However, such decisions do not approximate real-life social decision-making scenarios wherein we are more commonly making decisions about multiple familiar others. The reasons for these analytic decisions are similar to the ones for focusing on socially distant others. Some research questions are fundamentally concerned with how individuals make decisions affecting only one other person. Other questions might not easily accommodate the presence of more than one social agent or certain tasks (e.g., economic games designed to measure trust) are not easily adapted to include multiple individuals.

While pursuing research questions that only involve one target social agent is not inherently limiting, solely concentrating on these scenarios likely does the field a disservice. Humans do not uniquely congregate and interact in dyadic arrangements. Indeed, dyadic interactions are important, but everyday life decision-making scenarios are replete with even great complexity—anthropological work on decision-making indicates that multiple individuals are often affected by one's decisions (Henrich, 2002; Mathews, 1987). As such, only considering decision scenarios when one person is affected—essentially assuming decisions occur in a dyadic vacuum—does not capture the complexity of everyday life. It omits the possibility that preferences surrounding one social agent may be liable to change when additional contingencies are introduced. Indeed, individuals often have multiple social goals that are flexible and contingent upon multiple conditions (Pfeifer & Berkman, 2018). To assume that decision preferences regarding one individual are invariant across multiple decision contexts neglects much of what we know about motivation and how human resolve social goals (DeStasio et al., 2019) and thus fails to meaningful contribute to existing psychological theories (Berkman & Wilson, 2020).

Understanding how individuals make decisions that affect multiple individuals figures to boost the external validity of the social decision-making literature, in addition to serving other areas of psychology. For instance, recent work has suggested the personal relevance of health interventions affects the degree of their success (DeStasio et al., 2019; Hall et al., 2020; Yeager et al., 2018). This implies that if social decision-making research can enhance its external validity, its findings can be marshalled to help design and promote better health. For instance, by altering social decision experiments to include *multiple* close others as social agents and pitting outcomes for close others against one another, researchers may be able to estimate which close others are more powerful in swaying decision behavior and could therefore serve as integral components of intervention programs. Addressing the issue of involving multiple social agents in tandem with involving close others can also help psychologists better understand how individuals resolve competing social goals, potentially answering questions that have persisted since the early days of the field (Atkinson, 1957; Hull, 1931; Rotter, 1960). As I discuss below, a developmental approach may be helpful for answering these questions, as developmental differences in the importance of, and orientation towards, specific relationship types may provide exemplar populations to search how individuals resolve competing goals during social decision-making.

Target Population and Focal Social Decision-making Paradigm

Older adolescence and young adulthood present as a particularly pivotal developmental period for evaluating social decision making for two primary reasons. First, late adolescence is an extraordinarily unique phase in social development that renders it ideal for exploring the current phenomena (Crone & Fuligni, 2020). Children rely heavily on their parents for support and these relationships are frequently theorized to serve as anchors in children's social networks, whereas in adulthood evidence suggests individuals rely heavily on romantic partners for support and less so on platonic friends and parents (Hudson et al., 2015; Kahn et al., 2019; Ruhl et al., 2015; Theisen et al., 2018). That older adolescents rely heavily on both friends and parents, while also having more autonomy than younger adolescents to affect change in the lives of both these close others, means that individuals of this population serve as an ideal test case to understand how individuals resolve *strong* competing goals related to each close other (Crone &

Fuligni, 2020). Sampling individuals at other ages, by contrast, may not be as informative because it is unlikely that two types of relationships would simultaneously carry such high levels of importance. Relatedly, that both socioemotional processing and reward salience are most salient during the teenage years (Blakemore & Mills, 2014; Galvan et al., 2006) further bolsters the case for adolescence as a model test population because potential mechanisms underlying social decision preferences will likely be easier to statistically detect.

Another important reason to study adolescence is that disturbances to mental health and wellbeing tend to first onset at this time (Blakemore & Mills, 2014; Shulman et al., 2016). Crucially, evidence suggests that social decision-making and related psychological processes are often implicated in the same types of mental health and wellbeing disturbances that manifest during adolescence (Contreras-Huertas et al., 2020; Engelmann et al., 2015; Fernandez-Theoduloz et al., 2019; Gilbert, 2015; Harle et al., 2010; Lamba et al., 2020; Stamatis et al., 2020; Strang et al., 2017; Ting et al., 2021). Further, late adolescents in particular are experiencing newfound freedom and autonomy during this time, often having to make decisions within social contexts that are directly consequential to one's health (e.g., drinking, experimenting with drug use, sexual behavior; (Fromme et al., 2008; Shulman et al., 2016). Therefore, it is arguable that studying social decision-making during late adolescence will generate knowledge that will be helpful for understanding the developmental antecedents of mental health and wellbeing disturbances (those related to social decision-making, in this case). This would bode well for translational efforts, especially those geared at promoting continued wellbeing throughout development. In particular, knowledge gained from this study could potential information intervention efforts that depend on personalized, or precision, science about whom is most effective to deliver an intervention (e.g., delivering an intervention aimed at changing drinking behaviors, or combating loneliness among first year college students).

One final argument for studying this population rests on the assertion that adolescence is simply an intrinsically interesting and meaningful time for social decision-making research. Developmental science has shown that humans live in flux their entire lives—various features of their psychology continue to change from childhood through senescence (e.g., Jolly & Chang, 2019; Seaman et al., 2016). This would suggest trying to understand social decision-making from a generalist perspective is incomplete, or misleading at best. Plenty of work exists examining social decision-making in adults and, to a lesser extent, young children (Do et al., 2017; Sierksma & Shutts, 2020, 2021). By comparison, social decision-making is relatively understudied in adolescence. This is regrettable given the truly dynamic and unique nature of adolescence. This underscores the fact that basic science in both developmental psychology and social decision research could stand to benefit from the inclusion of more adolescent participants in social decision-making studies.

Although I surveyed social decision-making in scenarios involving prosocial behaviors and reciprocal exchanges above, I specifically chose to study social decision-making preferences in the context of risk taking for several reasons. Risk taking serves an ideal decision context because (i) it adequately captures the uncertainty associated with everyday choices, (ii) individuals encounter such scenarios with uncertainty in their outcomes on a regular basis, (iii) such scenarios not only accommodate outcomes for multiple affected agents but they also present an elegant yet powerful way to manipulate trade-offs for each agent, (iv) risky decisions are developmentally salient to my target population in a way that enhances the ecological relevance of my work (Steinberg et al., 2017).

Current Studies

This dissertation is comprised of three studies that will examine social decision-making under externally relevant contexts. Specifically, they attempted to understand how individuals in the midst of the transition from adolescence to young adulthood make decisions that involve conflicting outcomes for a parent and friend. In doing so, these studies characterized decision-making preferences involving multiple close others, helping determine whether individuals show reliable decision preferences for a particular close other or whether they simply focus on decision-level features while ignoring the identities of the implicated social agents. While the three proposed studies are to take place in a laboratory to ensure tight experimental control, that they involve making decisions about multiple close others is nevertheless a step towards enhanced external validity. The three studies here unpacked how individuals navigate social decision-making trade-offs involving two close others at three levels of analysis: behavioral, neural, and cognitive. The first study will establish a groundwork by documenting whether reliable preferences exist at the group level when making decisions regarding two close others. The latter two studies helped uncover potential mechanisms driving any potential preferences by examining how representations of these close others could potentially sway decision behavior. Study 2 assessed representations of parents and friends at the neural level, leveraging fMRI techniques to observe spontaneous representations of close others in ways that traditional methods have historically struggled to do so. Study 3 used natural language processing to extract cognitive representations of parents and friends from naturalistic written text data, in efforts to disambiguate the findings observed in Study 2.

Study 1: Estimating Social Decision-Making Preferences with Multiple Close Others at the Level of Behavior. Study 1 will use an experimental paradigm from behavioral economics to document whether individuals show preferences towards parents or friends—two important close others—during social decision-making. By pitting the outcomes for parents and friends against each other, I hope to incentivize individuals towards consistently making decisions that benefit a prioritized close other. Moderators will also be assessed in order to explain potential individual differences in decision preferences. Study 1 will thus be able to answer the basic question of whether decision behavior is modulated by the identify of affected social agents, as well as more complex questions about what features of relationships with social agents affect decision preferences.

Study 2: Probing Representations of Close Others at the Neural Level to Understand Social Decision Preferences. Study 2 will use fMRI to capture neural representations of two close others (a parent and friend) and evaluate the extent to which these representations are encoded as value-based signatures. In doing so, I aim to answer whether decision preferences among close others relate to how these close others are fundamentally represented in the brain in terms of value. This is notable for two reasons. First, prior imaging work related to social decision-making has largely demarcated neural ontologies of the decision maker's *generalized cognitive processes* (e.g., tolerance for ambiguity, aversion for guilt, etc.) and less so about how the decision maker represents *social targets* during decision-making. Second, work investigating neural representations of others and linking them to behavior have traditionally overlooked value-based processes. However, value signals are potent influences on human behavior and emerging evidence suggests that brain systems of valuation encode meaningful social information, such as social network position and behavioral traits (Braams & Crone, 2016; Hackel et al., 2015; Morelli et al., 2018; Parkinson et al., 2017; Zerubavel et al., 2015). While this work suggests value-based systems are important for encoding social information, it is virtually unknown whether the extent to which value-based signals permeate social representations of others (as opposed value-related regions discriminating between salient characteristics of others). The use of fMRI in this study is especially valuable because imaging approaches are helpful at uncovering latent psychological processes and states in ways that behavioral approaches cannot (Rhoads et al., 2021).

Study 3: Cognitive Representations and ties to Social Decision-Making. Study 3 will use written free-response data in conjunction with graph theory and natural language processing techniques to estimate semantic representations of close others. In turn, I will link structural features of these representations (operationalized as graph theoretic statistics computed over a network) to decision preferences. This is notable for reasons similar to Study 2. I am again examining mechanisms of social decision-making that involve the decision maker's cognitions about the *social targets* instead of intrinsic cognitions about decision-level features (e.g., aversion to risk). This time I am linking semantic and conceptual knowledge the of social agent to decision preferences, which has not yet been attempted in the literature.

Study 1: Social Decision-Making with Close Others in a Risky Context

As social creatures, humans must often consider how their decisions will impact multiple close others. However, as noted above, psychologists rarely study social decision-making as it relates to close others, let alone more than one close other at once. It is thus unclear how individuals navigate decision scenarios involving trade-offs involving those closest to them. In the current study, I will force individuals to make decisions that pit financial outcomes for a parent and friend against one another in hopes of better understanding social decision-making preferences when close others are at stake. I aim to conduct this research in older adolescents transitioning to young adulthood because (i) this is the transitory phase where preferences, identities and behaviors begin to crystallize (Damian et al., 2019; Roberts et al., 2003; Robins et al., 2001), (ii) older adolescents still display peer-oriented characteristics of adolescence yet are still highly dependent on parents (Silva et al., 2016) while being (iii) simultaneously required to adapt to novel roles and adultlike situations (Arnett, 2014; Fuligni, 2018). Because relatively little work exists that attempts to pinpoint mechanisms behind social decision-making preferences, I begin to address this issue by considering the role of key moderating factors, such as age and relationship quality with parents and friends. This study will document whether individuals have clear and consistent decision preferences when close others are affected. This ultimately matters because real life decisions made by older teens transitioning to young adulthood frequently impacts both friends and family.

Late Adolescence is Marked by a Unique Social Ecology. Psychologists, psychiatrists, and physicians alike have attempted to characterize the boundary that demarcates adolescence from young adulthood for several decades (Curtis, 2015; Gould, 1972; Greenberger, 1984; Levinson, 1986; Sawyer et al., 2018). Though considerable debate exists over whether adolescence ends at age eighteen, twenty, or even twenty-four, most scholars generally agree that late adolescence—and the bridge to young adulthood—is a unique transitory period unlike any other in the lifespan. During this period, individuals face newfound contexts and responsibilities and a seemingly ever-shifting social milieu (Arnett, 2000, 2014; Fromme et al., 2008; Roberts et al., 2003). At the center of these changes lie the relationship dynamics between individuals and their parents and peers. Individuals navigating late adolescence are oriented towards their peers but also depend on their parents for financial, emotional, and social support (Arnett, 2014).

Peers, especially close friends, affect the way late adolescents construe their identities (Hopmeyer & Medovoy, 2017; Pfeifer et al., 2009; Roisman et al., 2004), develop opinions

(Welborn et al., 2015), and make decisions under conditions of uncertainty (Fromme et al., 2008; Riedijk & Harakeh, 2017; Silva et al., 2016; White et al., 2006). Specifically, adolescents' propensity to alter their cognitive, behavioral, and socioemotional processes (like those mentioned above) appears to be driven over explicit concern over the opinions and preferences of their peers (Pfeifer & Berkman, 2018; Steinberg & Morris, 2001). This concern appears to stem from desire to gain social status, acceptance among same-age peers, and to facilitate exploration of a novel social milieu (Albarello et al., 2018; Blakemore, 2018; Brown et al., 1993; Spear, 2000b, 2000a). Despite the seemingly monolithic importance of friends, parents also remain important in the lives of late adolescents. Like friends, parent-adolescent relationships also impact real-world decision-making (Abaied & Emond, 2013; Carlson, 2014; Guassi Moreira & Telzer, 2018b), refine existing identities (Kaniušonytė & Žukauskienė, 2018), and promote wellbeing (Lucas-Thompson, 2014; Needham, 2008). Similar to peers, adolescents also appear to care enough about the opinions and preferences of their parents to alter cognitive, behavioral, and socioemotional processes, with the added. In the case of parents, this concern often appears to be driven by a sense of obligation, and desire to contribute, to the familial unit.

As individuals transition progress through adolescence, parents appear to become *more* central. Older adolescents, continuing into young adulthood, report higher levels of cohesiveness and obligation towards parents while becoming increasingly receptive to their advice (Carlson, 2014; Tsai, Telzer, & Fuligni, 2013). This illustrates how parent-child relationships are an important source of stability for late adolescents at a time when their social environments are becoming increasingly unstable (graduating from high school, attending college or entering a dynamic workforce, etc.). These pieces of evidence demonstrate that both parents and peers play crucial roles in the lives of young adults but it is unknown how these close others are prioritized when pitted against one another.

Age and Relationship Quality as Moderators. Evidence suggests that age and relationship quality are likely to influence how older adolescents make social decisions involving parents and friends (Almas et al., 2010; Hackel & Zaki, 2018; Huck et al., 2012; Powers et al., 2018). With respect to age, I anticipate that older individuals will favor parents over friends for a number of reasons. First, parent-child relationship quality increases during late adolescence and through young adulthood (Tsai et al., 2013), coinciding with high rates of friendship turnover (Arnett, 2014). Second, parent-child relationships are often characterized by high financial investment from parents. Because individuals tend to view acts of generosity such as these favorably (Hackel & Zaki, 2018), older individuals may be more likely to favor their parents because more cumulative financial investment has occurred. With respect to relationship quality, I expect that greater relationship quality with a given close other will be related with a greater tendency to prioritize them during decision-making. Relationship quality has been shown to affect decision-making behavior—and its cognitive antecedents—in other domains (Guassi Moreira & Telzer, 2018a, 2018b; Mattanah et al., 2011).

Current Study. I investigated how older adolescents make decisions that have opposing consequences for a parent versus a friend. By pitting parent and friend consequences against each other, I aimed to force individuals to reveal their social decision preferences, as they involve close other, in a context that emulates real-world trade-offs. Participants completed one round of a decision-making task in which they could either reward a parent at the expense of a friend, or forgo the reward for a parent to avoid a loss for a friend. A second, counterbalanced round featured opposite tradeoffs (friend reward at expense of parent, etc.). This manipulation is theoretically well suited to address how individuals make decisions about parents versus friends because it is *incentive compatible*. In other words, the task is inherently designed to motivate participants to reveal their degree of preference between a parent and friend (strong parent, strong friend, or roughly equivocal).

Hypotheses. My primary hypothesis contained a set of competing predictions. There is ample evidence to suggest that parents *and* peers each play crucial roles in the social ecology of young adulthood. If parents are more central than peers, participants would be more inclined to make decisions that benefit a parent at the expense of a friend. However, if peers are more important than parents, then individuals would be more likely to make decisions that benefit a friend at the expense of a parent. I note here that because extant evidence highlights how both parents and peers exert strong pulls over behavior, it is appropriate to formulate a competing hypothesis such as this one.

My secondary hypotheses are that relationship quality will predict social decision preferences (e.g., better relationship quality with one's parent will be associated with a tendency to favor a parent, better quality with friend will be associated with a tendency to favor a friend), and that older participants, compared with younger participants, will be more likely to make decisions that benefit a parent at the expense of a friend

Method

Participants. Participants for this first study were comprised of 180 late adolescents and young adults from the West Los Angeles area in the United States. Considerations regarding sample size justification are enumerated below in a separate section. Six participants were excluded from analyses (4 due to noncompliance with experimenter instructions, 2 due to technical errors with data collection equipment). The final sample included 174 late adolescents (Age: mean = 20.72 years, SD = 2.16, range = 18.06 – 30.81; 54 males). The considerable ethnic diversity of Los Angeles was reflected in the sample, as anticipated, helping enhance the generalizability of the proposed study (Simons et al., 2017). Racially, 48% participants identified as Asian, 29% as Caucasian, 2% as African American, 6% as mixed race, 1% as Native American/Alaskan Native, and 10% as 'other race'. Ethnically, 18% of participants identified as Latinx. Recruitment practices included using the University of California, Los Angeles undergraduate psychology subject pool, posting recruitment materials at large in the West Los Angeles community (Westwood, Santa Monica), and disseminating recruitment materials at a local community college through faculty (Santa Monica College); 92% of participants were enrolled in classes as UCLA. Participants were compensated with course credit or a $20 (USD) cash payment, depending on method of recruitment. Because a secondary question in this particular study involved examining age effects, I recruited individuals between the ages of 18 and 30 years. I decided on an upper bound of 30 years since that age is unambiguously not included in the many contemporary definitions of adolescence. All participants provided written consent in accordance with the policies of the UCLA Institutional Review Board.

Sample Size Considerations. Several variables factored into sample size determination. I considered sample sizes from prior studies, anticipated effect sizes, the nature of the proposed statistical models (described in a following section), in addition to the availability of tools to estimate power. During this consideration, it became apparent that conducting a formal power calculation would be onerously difficult. As detailed in a subsequent section ('Bayesian Multilevel Logistic Regression' subsection of the 'Analysis Plan'), this study uses Bayesian multilevel logistic regression models. While these models are invaluable statistical tools that provide several advantages over traditional approaches, they are accompanied by the following challenges for power calculations. (i) There are no straightforward analytic methods to estimate power; (ii) key statistical information (e.g., posterior summary statistics) needed to conduct custom simulations is rarely, if ever, reported in prior work using similar analysis methods; (iii) effect size metrics for logistic regression models are not well established and guidelines for extant effect sizes are unclear (Chen et al., 2010). Although it would be difficult to conduct an effective power analysis, it is nevertheless important to recognize that statistical power is a critical part of study design. To adequately address this issue, I took the following measures. First, I set my target sample size to be much larger than recently published similar investigations of adolescent decision-making and social influence on adolescent behavior (Botdorf, Rosenbaum, Patrianakos, Steinberg, & Chein, 2016, *N* = 104; Guassi Moreira & Telzer, 2018, *N* = 63; Riedijk & Harakeh, 2017, *N* = 63; Silva, Chein, & Steinberg, 2016, *N* = 100 per cell; Silva, Chein, & Steinberg, 2020, *N* = 45 per cell; Van Hoorn, Crone, & Van Leijenhorst, 2017, *N* = 73). Second, the key manipulation occurred within subjects to increase power (Scherbaum & Ferreter, 2009). Because each individual will provide between 48 and 744 individual choices (contingent on the number risky decisions made on each round), this fact should help enhance confidence that the proposed study is adequately powered.

Experimental Protocol

Overview. The general experimental protocol that was performed is outlined here. Participants were be greeted by an experimenter upon arrival to the lab and gave informed consent. Afterwards, they were asked to nominate a parent and close friend of their choice and complete a salience procedure. Participants then answered a series of self-report questionnaires and completed the computerized experimental paradigm. Each section of the protocol, beginning with the salience procedure, is described below in greater detail. I conclude this *Experimental Protocol* section with a brief description of the types of rewards (real vs hypothetical) participants could earn for their parent and friend during the task.

Parent-Friend Salience Procedure. During the consent process, participants were informed they would be required to make decisions on behalf of a parent and close friend of their choice and were be asked to choose one such individual. Because these close others were not be accompanying participants to study, I reasoned it was important to heighten the salience of such decisions for participants (Shah, 2003). In order to achieve this, I asked participants to complete a set of prompts on a physical sheet of paper. The prompts asked participants to provide basic information about each close other (e.g., name, age, sex), briefly recount a memory they have with each person, and list a handful of words and phrases describing each person.

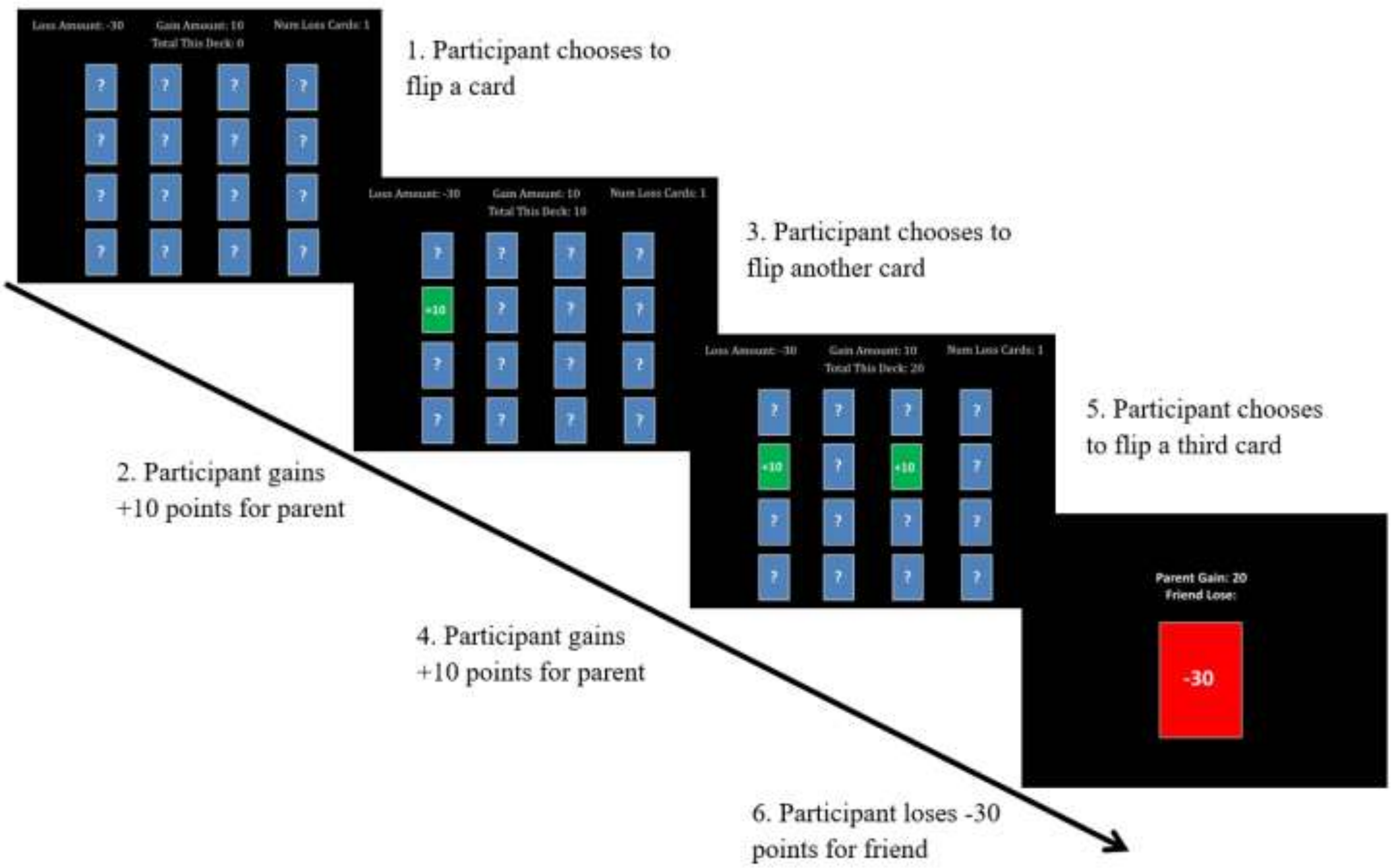
Self-Report Measures. Following the salience procedure, participants completed a series of self-report measures on a laboratory computer via Qualtrics (an online survey administration platform). The key measure of interest was the Inventory of Parent and Peer Attachment (IPPA; Armsden & Greenberg, 1987), a frequently used measure of subjective relationship quality (Branje et al., 2010; Fanti et al., 2008). This measure was administered to test the hypothesis about the association between individual differences in social decision preferences and relationship quality. The IPPA is comprised of 28 items about parent relationships and 25 items about friend relationships across three domains (trust, communication, and alienation) and is implemented using a five-point Likert scale (1 = *almost never or never,* 5 = *almost always or always*). The original instrument generically asks about parents and friends, but I adapted it to create a minimally-modified version that instructs participants to complete the instrument their nominated parent and friend. Sample items include “When parent respects my feelings” and

“When I discuss things, my friend considers my point of view”. Responses were reverse scored where appropriate and averaged to yield a single mean score for parent and friend relationship quality, respectively. Though the measure has historically been shown to possess good reliability, recent work in psychometrics has suggested researchers should assess reliability using model-based techniques (Rodriguez et al., 2016b, 2016a). In light of this, I calculated coefficients ω and ω-hierarchical as indices of reliability as opposed to Cronbach’s alpha. The measure demonstrated good reliability (parent ω = 0.96, parent ω-hierarchical = 0.76; friend ω = 0.94, friend ω-hierarchical = 0.81) Additional self-report questionnaires tapped domain-specific risk-taking, sensation seeking, and family obligation.

Social Decision-making Paradigm. Here I introduce the proposed experimental paradigm. Because it involves modifications to an existing task, I first detail the traditional task parameters and then describe the proposed changes that I argue allowed the task to capture social decision-making preferences.

I used a modified version of the computerized “hot” Columbia Card Task (CCT) to assess social decision-making preferences involving conflicting outcomes for parents and friends (Figner et al., 2009; van Duijvenvoorde et al., 2015). The hot CCT is a widely used experimental paradigm that measures risky decision-making in an incremental, stepwise manner. During an experimental session, participants completed two runs of the CCT. One run consisted of 24 rounds (sometimes referred to as ‘game rounds’); each round comprised a series of iterative decisions, ranging from as few as one to as many as sixteen. During each round, participants were shown a set of sixteen overturned cards (i.e., collectively, a deck). They were told each card is associated with either a gain or a loss of points, and that the objective of the task is win to points by iteratively turning over cards. Participants were notified of a header above each deck (i.e., the set of 16 cards) that provided information about the overturned cards. The header indicated (a) the total number of loss cards in the deck (one or two), (b) the point value of each loss card (-30 or -60), (c) the point value of each gain card (10 or 20), and (d) a running total that tracked the points they are earning for that deck (Figure 1.1). These three possibilities (number of loss cards, loss card amount, gain card amount) were crossed to yield eight distinct deck types, each of which was presented three times during a run (hence 24 rounds).

Figure 1.1. An example of the modified Columbia Card Task (CCT).



Note. Here the participant is shown winning 20 points for their parent at the expense of losing 30 points for a friend.

Each round began with a score of zero points and all cards overturned. Participants were required to choose between turning over a card—a risky choice, given that the outcome could elicit a gain or loss—or not turning over a card. Electing to not turn is termed ‘passing’, and is considered a ‘safe’, non-risky choice. If participants chose to turn over a card, the computer randomly selected a card and turned it over. Choosing to pass, by contrast, ended the round and participants could not gain or lose any additional points (passing is akin to ‘cashing out’ at a casino). Each round lasted until the participant decided to pass or randomly flipped a loss card. Participants were informed the computer selected cards to flip at random. In reality, the first three risky choices for any deck were always rigged as yield a gain card, helping ensure participants did not lose too early and feel disproportionately discouraged from taking further risks. Participants completed four practice rounds to ensure proper understanding of the task. A trained experimenter did not allow them to proceed unless they demonstrated a clear understanding of the rules.

As previously mentioned, the CCT was modified to assess late adolescents’ social decision-making preferences between parents and friends. During one run of the task, participants were informed all points associated with gain cards would be awarded to their nominated parent, whereas any losses associated with each loss card would be incurred by their nominated friend. The opposite was true during the second run (gains solely benefit friend, losses solely incurred by parent). The run order of the two conditions (Parent Gain-Friend Lose, Friend Gain-Parent Lose) was counterbalanced between subjects to ensure ordering of conditions did not affect decision behavior. Critically, I argue this manipulation models real-world trade-offs. Individuals were forced to make decisions that benefitted a close other at the potential expense of a second close other—there is never a trial in which only one close other is affected, ensuring there is always a potential cost for favoring one close other. Relatedly, this feature of the experimental design also ensure it is incentive compatible. For example, during a Parent Gain-

Friend Lose run, flipping over a card to earn points for a parent necessarily means risking a friend's points. Alternately, protecting a friend from a loss by passing would mean that one is giving up the chance to win points for a parent. This means participants were always incentivized to reveal their social decision preference. Text describing the condition of the current run ('Parent Gain | Friend Lose' or 'Friend Gain | Parent Lose') was presented at the bottom of the screen on each trial as a reminder to participants. Outcomes were also clearly labeled to ensure participants understood what each close other may have won or lost following a given set of cards. The task was programmed and administered using the open-source, python-based PsychoPy software (Peirce, 2007). An experimenter remained present and unobtrusively monitor the participant during completion of the task in order to ensure participant focus and diligence.

Manipulating Reward Type: Real and Hypothetical. Here I briefly consider the issue of using hypothetical and real rewards in my social decision-making paradigm, and justify my solution to the issue. Extant evidence offers conflicting narratives on whether to use real or hypothetical rewards when conducting decision-making experiments with monetary consequences. Ideally, real rewards would always be used in social decision experiments that involve monetary outcomes. Yet this is not practical for a number of reasons (e.g., fiscal research costs, ethical considerations with financially vulnerable populations, etc.). Therefore, researchers have often resorted to using hypothetical rewards in lieu of real monetary payouts. Historically, evidence from behavioral economics shows there are no differences in choice behavior when using real or hypothetical rewards (Locey et al., 2011). However, these studies were conducted under a narrow set of circumstances: participants made self-oriented decisions (i.e., participants made choices only affecting themselves), a specific set of computerized paradigms were used (most typically discounting), and they are oft divorced from any kind of social context whatsoever. Social psychology and neuroeconomics suggests the distinction between real and hypothetical rewards may matter: the two reward types recruit different neural systems (Bray et al., 2010; Kang et al., 2011; Miyapuram et al., 2012; Scholl et al., 2015), social contexts affect how the two reward types are perceived (Pronin et al., 2008), and different types of decisions (e.g., moral decisions) appear to elicit differential behavior between the two types of rewards (FeldmanHall et al., 2012).

Although the existing evidence on this matter is hardly unequivocal—and the issue deserves additional empirical scrutiny in its own right—the presence of conflicting evidence suggests there is a chance that reward type would differentially impact results in the current dissertation. This design choice therefore required careful consideration. In light of this, I decided to include both reward types in this study as a between subjects manipulation. Half of the sample was instructed to play as if the points earned during the task could be redeemed for tangible material goods or services. Specifically, they were told to complete the task as if their parent and friend were given an initial endowment of points that could be redeemed for tangible rewards, and that their decisions could affect the final total of their parent's and friend's endowments. Although it was emphasized that participants are to play as if rewards were not hypothetical, it was made clear that their parent and friend will not actually receive any actual rewards. The other half of the sample was given the same instructions, except with real rewards. Participants in this half of the sample were required to provide a mailing address for their parent and friend, and were informed that their parent and friend each began with an endowment of points worth $5 which could be affected by their decisions during the experiment. Participants were shown envelopes with their parent's and friend's respective address containing $5. To help ensure these participants believed the rewards were not fixed, the experimenter informed them a computer script would be run immediately following the session to determine the amount or lost for parents and friends. Afterwards, the experimenter ran a dummy script to give the appearance that actual calculations were being made. Participants were then be notified they won an additional $2 for their friend and an additional $1 for their parent (that these values are fixed was unbeknownst to the participants). The earnings were mailed after the session, and participants were debriefed about the fixed reward amounts following the session.

Analysis Plan

I describe my analysis plan by first outlining a parameterization of the CCT via a risk-return decomposition, and then explaining how that parameterization was used in Bayesian multilevel logistic regression analyses. I then briefly review methods for selecting prior probability distributions for such analyses before justifying the priors I selected herein. Finally, I conclude with a section on how I elected to perform inference in this Bayesian context.

CCT Parameterization via Risk-Return Decomposition. Because different facets of risky scenarios tend to exert unique influences on decision-making (van Duijvenvoorde et al., 2015), it can be advantageous for statistical modeling to decompose risky scenarios into their constituent components (Richards et al., 2013; Tobler et al., 2009). One such method is known as a risk-return decomposition. Popular in behavioral economics, this parameterization involves explicitly partitioning risky scenarios into a pure return component (the expected value or, anticipated reward, of taking a risk) and a pure risk component (the outcome variability associated with the risky decision alternative). This decomposition thus dissociates the independent components of return and risk for every trial. Using these components in a statistical model has a few advantages. First, it can serve as a manipulation check to ensure participants are properly completing the task (e.g., if participants are not more likely to flip over a card as return increases, it may be indicative of mischievous, unfocused, or confused responding). Second, it enriches knowledge of social decision-making by facilitating future comparisons of return and risk sensitivities in social and non-social decision-making contexts. Finally, it gives one the flexibility to test for moderating influences on social decision preference (e.g., perhaps the effect of risk is more salient in one condition compared to the other).

Consistent with recent work using the CCT (van Duijvenvoorde et al., 2015), return was operationalized as the expected value (EV) of a decision to flip over a card:

$$\text{EV (Return)} = (\text{Gain Probability x Gain Amount}) + (\text{Loss Probability x Loss Amount})$$

Risk was operationalized as the standard deviation of the distribution of all possible outcomes associated with a decision to turn over a card:

$$\text{SD (Risk)} = \sqrt{\text{Gain Probability x (Gain Amount} - \text{EV})^2 + \text{Loss Probability x (Loss Amount} - \text{EV})^2}$$

Because the decision to pass did not result in the gain or loss of points, the EV and SD for choosing to pass were both zero. Given the combination of parameters for each possible game round (number of loss cards, value of loss cards, value of gain cards), the EV values ranged from -37.37 to 16.88 and SD values ranged from 9.68 to 40.00[3], with each trial within each deck receiving its own unique EV and SD values (relative to other trials in that same deck).

Bayesian Multilevel Logistic Regression.

The goal of this study was to understand whether individuals evince consistent social decision-making preferences for a parent or a friend in a risky context, while controlling for

[3] Technically, zero is the lowest possible SD value, provided one has flipped all of the gain cards for a given round and only loss cards remained. Because this event is unlikely, I include here the next lowest possible value to give readers a realistic range of what participants will encounter. Similarly, -60 is technically the lowest possible EV, but a more realistic EV is -37.37.

other salient, low-level features of a given decision (i.e., risk and return values) and other confounds (age, sex). There was also a secondary emphasis placed on understanding moderators of social decision preferences, namely relationship quality with one's parent and friend, in addition to the type of reward at stake (hypothetical v real). This means this analysis required an analytic framework that could (i) handle a binary outcome variable, (ii) accommodate within- and between-person predictors, and (iii) properly account for the hierarchical nature of the data (i.e., decisions nested within individuals). These three requirements suggested some kind of multilevel (i.e., hierarchical or random coefficient) logistic regression was necessary. A multilevel framework could adequately handle the nesting structure of the data, whereas the logistic component of the model could handle the binary dependent variable[4].

Rationale. I implemented this analysis in a Bayesian framework using the *brms* package (Bürkner, 2017) in the R statistical software (R Core Team, 2017). The Bayesian approach provides several advantages over traditional frequentist methods that I enumerate below.

First, as the output of a Bayesian analysis is a distribution, rather than a point estimate, statistical inference can be performed in a graded manner with greater nuance (Kruschke, 2013). Specifically, because the posterior probability distribution represents the likelihood of possible parameter values given the data, one can analyze where the bulk of the posterior's mass falls in relation to a null value or region and conclude whether there is evidence for (i) a robust effect, (iii) a null effect, or (iii) an effect that rules out a specific sign (e.g., a regression slope whose sign is positive or zero, but not negative). Second, Bayesian modeling offers a principled way to conduct regularization of model parameters (Simpson et al., 2017). Regularization is the

[4] It did occur to me the data could be marshalled to fit into a more traditional framework (e.g., t-test, ANOVA) via aggregation (summing the number of choices within one more conditions). However, this approach comes with the disadvantage of disregarding trial-by-trial variability and limiting precision of the analysis.

statistical practice of systematically and modestly biasing parameter estimates in order to reduce the sample-to-sample variability of such estimates (McNeish, 2015). Doing so means statistical model parameters are less likely to be influenced by the random noise of a given dataset and should theoretically generalize better when applied to new data. In a Bayesian context, this is achieved by incorporating uninformative (flat) or weakly informative (diffuse) priors into the model, as discussed in detail below. Third, Bayesian approaches inherently facilitate cumulative and incremental science as outputs of old analyses can be directly incorporated into new analyses, such as using a posterior distribution from a previous analysis as a prior distribution for a new analysis. Finally, because a Bayesian analysis (i) yields a posterior distribution as its outcome and (ii) can estimate hundreds of parameters, it is suitable to model psychological phenomena in such a way that captures their complexity while also providing a generative account of such phenomena (i.e., model parameters can be used to generate new data that resembles observed data). While this practice (generative modeling) is not directly employed in this dissertation, I find it helpful to select an analysis that is amenable to this for sake of facilitating future research (whether it be follow-up investigations of my own, or those conducted by other scholars interested in the same research questions).

I specifically chose to use the *brms* package because it is built upon the Stan platform for statistical modeling and computation (Stan Development Team, 2021). While Stan is similar to other Bayesian statistical software in that it implements Markov chain Monte Carlo algorithms to draw samples from the posterior distribution, it is unique from others insofar that it is the most flexible and effective at doing so. This is because Stan uses a particular type of MCMC sampling, Hamiltonian Monte Carlo (Duane et al., 1987; Neal, 2011), and its extension, the No-U-Turn Sampler (NUTS) (Hoffman & Gelman, 2014). The main advantage of these combined algorithms is their reduced convergence time and an effectively wider range of prior distributions offered for user specification (i.e., conjugacy is not required to improve sampling speed). Further technical descriptions of this the package can be accessed in user guides for Stan (https://mc-stan.org/docs/2_27/stan-users-guide/index.html) and *brms* (https://www.rdocumentation.org/packages/brms/versions/2.15.0).

Model. Decisions on the *i*-th trial from the *j*-th participant were modeled as being distributed Bernoulli.

$$\text{Decision}_{ij} \sim \text{Ber}(p_{ij}) \tag{1}$$

Equation (1) also represents the likelihood of the model (i.e., the probability of observing the current data distributed over a range of parameter estimates). The Bernoulli distribution is frequently used to model binary outcomes, and takes a single parameter (p) describing the probability of ‘success’. Here, p_{ij} represents the probability of the *j*-th participant making a risky decision (i.e., turning over a card) on the *i*-th trial. The log odds of these probabilities were modeled as a linear combination of trial-level variables; an intercept (b_{0j}), the experimental condition (b_{1j}; 1 = Parent Gain-Friend Lose, 0 = Friend Gain-Parent lose), return (b_{2j}; EV), and risk (b_{3j}; SD).

$$\ln\left(\frac{p_{ij}}{1 - p_{ij}}\right) = b_{0j} + b_{1j}\text{Condition}_{ij} + b_{2j}\text{Return}_{ij} + b_{3j}\text{Risk}_{ij} \tag{2}$$

Critically, b_{1j} is the key parameter of interest, as it encodes social decision preferences. A value equal to zero indicates no preference, a positive value indicates a parent-over-friend preference, and a negative value indicates a friend-over-friend preference. Because this is a *logistic* regression, it is important to note that coefficients represent expected changes in logit units – that is, a one unit increase in any predictor will be associated with an expected change in the log odds of a risky decision equivalent to b (referring to a generic coefficient). Logit units can be converted to an odds ratio (i.e., the expected change in the odds) by exponentiating a given coefficient (i.e., exp(b)).

Notably, coefficients associated with these trial-level variables can be decomposed into population-level (γ) and group-level (u) parameters. Population-level and group level parameters are loosely analogous to the notions of fixed and random effects, respectively, in the frequentist framework. Between-subject variables were included in the model, as moderators of the effect of condition (social decision preferences). This means the slopes associated with the intercept (b_{0j}) and condition (b_{1j}) were parameterized as equations (3-4).

$$b_{0j} = \gamma_{00} + \gamma_{01}PRQ_{ij} + \gamma_{02}FRQ_{ij} + \gamma_{03}Age_{ij} + \gamma_{04}Sex_{ij} + \gamma_{05}RewardType_{ij} + u_{oj} \qquad (3)$$

$$b_{1j} = \gamma_{10} + \gamma_{11}PRQ_{ij} + \gamma_{12}FRQ_{ij} + \gamma_{13}Age_{ij} + \gamma_{14}Sex_{ij} + \gamma_{15}RewardType_{ij} + u_{1j} \qquad (4)$$

Here, PRQ refers to parent relationship quality (continuous), FRQ refers to friend relationship quality (continuous), age was measured in years (continuous), sex reflected biological sex (0 = male, 1 = female), and reward type referred to whether rewards were real or simulated (0 = simulated; 1 = real). Between-subjects predictors were not added to the slopes for return (b_{2j}) and risk (b_{3j}), as notated in equations (5-6).

$$b_{2j} = \gamma_{20} + u_{2j} \qquad (5)$$

$$b_{3j} = \gamma_{30} + u_{3j} \qquad (6)$$

Last, I note that I built up to this full model in a series of four steps. The first model just involved within-subject (i.e., trial-level) predictors to understand main effects (Model 1). Next, I added parent and friend relationship quality as moderators of the effect of condition to determine whether social decision preferences depended on relationship quality (Model 2). Afterwards, I added age, sex, and reward type as main effects to rule out potential confounding effects (Model 3). Finally, I took this procedure a step further by running the full model described above, ensuring that any influences of relationship on condition could not simply be explained by age or sex (Model 4). This model also evaluates whether reward type plays a meaningful role in shaping social decision preferences. Models were fit using the default sampling procedures in the *brms* package (no thinning, 4 chains, 4000 samples per chains, 2000 discarded warm-up samples).

Prior Selection and Justification.

Bayesian analyses require the specification of a prior probability distribution (or simply a 'prior'). Plainly, a prior is a vehicle for researchers to incorporate subject matter expertise into their analyses by specifying their beliefs about the probability of possible parameter estimates over a range of candidate values (in the form of a distribution). This is in contrast to frequentist approaches using maximum likelihood estimation (MLE), which do not allow room for subjective expertise to influence the model because they only estimate the probability of the observed data over a range of candidate parameter estimates (known as the likelihood). A Bayesian analysis takes advantage of both quantities by combining priors and likelihoods to yield the posterior distribution (a distribution of possible parameter values, given the data). More precisely, the prior distribution is used to weight the likelihood, thereby 'giving room' for subjective expertise to influence an analysis. Illustrating this process anecdotally, if I clumsily stumbled off a curb and felt knee pain, I would probably reach the conclusion I had a slight hyperextension rather than a torn ligament because my previous physical experiences with pain and injuries help downweight catastrophic conclusions and lend confidence that something less severe occurred[5] (i.e., it's unlikely for one to tear a ligament when falling off a curb). An alternative way of conceptualizing this process is by what Kruschke refers to as 'reallocation of credibility' (2011). In this conceptualization, the prior represents a researcher's belief about a given research question before data collection. Once data are collected and used to calculate the likelihood, the prior's credibility among parameter estimates is shifted according to what the likelihood deems to be more or less likely.

The incorporation of priors can be both a blessing and curse. Priors represent powerful tools if applied appropriately, but can result in nonsensical, incorrect, or misleading inferences when applied inappropriately. Statisticians have spent decades formulating guidelines and recommendations for prior specification. While there are truly no 'hard and fast' rules about which specific priors to use—every research question is different and what is appropriate in one context may be wildly inappropriate in another (Gelman et al., 2017)—there are plenty of formal guidelines. For this dissertation, I considered two types of priors: non-informative (i.e., vague) and weakly informative. Non-informative priors are typically uniform or have large enough scales so that they strongly resemble or approach uniformity (e.g., $N(0, 10e6)$). Weakly informative priors impose more influence on estimates, but not in any way that would drastically sway estimates to favor the prior in the presence of otherwise compelling evidence (e.g., $N(0, 0.75)$ for a cohen's d effect size in behavioral science research). In other words, non-informative priors assume all parameter values are equally likely, whereas weakly informative priors are

[5] I acknowledge that the likelihood already accounts for the fact that mild knee pain following a stumble is relatively more consistent with hyperextension than a torn ligament, but the point here is to show how the prior helps modulate the credibility placed on certain conclusions (or parameter estimates, hypotheses, etc.). In other words, the prior helps 'nudge' the likelihood towards a more realistic conclusion via the use of subjective expertise.

modestly confident that parameter values fall in a particular range[6]. These categories are in contrast to informative priors, which represent highly specific beliefs about model parameters (e.g., $N(0.4, 0.2)$). I did not consider informative priors in the current dissertation for reasons outlined in my deliberation between non-informative and weakly informative priors (below).

In selecting a prior for this study, I was hoping to achieve two goals. First, I was aiming to avoid adding substantial bias to the analysis. I was particularly concerned with avoiding a prior that would bias the results in revealing an effect that was not actually present (e.g., suggesting a clear social decision preference when one did not actually exist). Second, I was hoping to achieve principled regularization of model parameters (briefly described above). While scientists traditionally use and discuss regularization in the context of predictive modeling (Hine & Usynin, 2005; Stiglic et al., 2015; Xiao et al., 2018; Yao & Yang, 2016), regularized approaches can also aid inference insofar that they minimize the influence of noise in parameter estimation (Efron & Morris, 1975, 1977; James & Stein, 1992). Thus, regularization can be considered a tool to enhance generalizability of model parameters. Together, these two goals meant I needed to select a prior that did not exert excessive or undue influence on model parameters but could also regularize estimates and slightly bias them towards zero. This led me to priors in the categories of non-informative and weakly informative.

When considering between these two categories, it was important to keep in mind what each prior conceptually represented. Non-informative, (i.e., flat, uniform) priors assign every possible parameter value the same density. This feature renders non-informative priors quite popular, as many researchers often look to strictly limit the influence of the prior on results. The

[6] Mind you, the range for weakly informative priors is still intended to be quite wide. The normal prior from the Cohen's *d* example given above essentially encompasses all the realistic effect estimates in psychology (95% of parameter values falling between -1.5 and 1.5).

flat, non-informative prior seemingly accomplishes this goal, as it does not influence the central tendency measures of the posterior distribution. However, non-informative priors come with two notable drawbacks. First, unless one has good reason to restrict the range of these priors, they will not appropriately regularize parameter estimates because they place equal mass across the entire range of possible parameter values. There is thus no small-yet-systematic biasing of the coefficients and subsequently no regularization. Second, and more concerning, is that if one's data are sparse enough for certain (often extreme) values of the parameter space, the non-informative prior actually becomes misleadingly informative. This is because non-informative priors tend to place an inappropriate amount of mass on values that are highly implausible (e.g., a coefficient of 100 in a logistic regression setting) in the absence of data that speak to the likelihood of these values. The drawbacks of the non-informative prior stand in stark contrast to the weakly informative prior, which can accomplish the opposite: diffusely pooling the mass of the prior distribution about a single point (typically a null value such as zero) has the desired regularizing effects while not giving weight to extreme and implausible values.

Given the discussion in the above paragraph and the previously discussed goals for the analysis, it was apparent that weakly informative priors were most suitable for this analysis. Thus, all fixed effects received a standard normal prior ($N(0,1)$). The normal distribution was selected because I did not have strong beliefs of asymmetry in the parameter space, and I did not have a reason to believe that fatter tails (i.e., as seen in a t-distribution) were necessary given the logistic regression model. The location and scale (mean, standard deviation) parameters of 0 and 1 were selected because (i) zero corresponds with the null value and regularization typically occurs by biasing coefficients to a null value, and (ii) a standard deviation of 1—in logistic regression—would virtually cover the entire range of plausible parameter estimates (effects greater than |3| in logistic regression correspond to enormous effect sizes on an odds scale, certainly larger than would be expected in behavioral science research). The random effects from the model were drawn from a student's distribution ($t(3, 0, 2.5)$). This t-distribution was used at the recommendation of the 'brms' developer, who notes that group-level effects often need require distributions with fatter tails.

Inference Criterion. I performed inference on the posterior samples by using the region of practical equivalence method popularized by Kruschke (2011, 2013). This method involves three steps.

First, a credible interval (CI)—a span of the posterior distribution capturing a user-defined portion of its mass—is computed for a given posterior. For this study—and those that follow—I use 89% credible intervals[7] upon the recommendation that wider intervals (e.g., 95%) are more to sensitive Monte Carlo sampling error (Makowski et al., 2019; McElreath, 2018). All CIs specified in each study were computed using the Highest Density Interval (HDI) method (*bayestestR* package; Makowski et al., 2019), which ensures that all values within the interval have a greater density than those outside it.

The second step requires the specification of a region of practical equivalence (ROPE). A ROPE is a user-defined interval in the parameter space whose values are deemed virtually equivalent to a null value (e.g., zero). In other words, the ROPE spans effects of such little magnitude that they are, for practical purposes, considered comparable to the null value.

Finally, the degree of overlap between the CI and ROPE is inspected and compared to the inferential criteria specified by Spiegelhalter and colleagues (1994), allowing one to perform

[7] 89 is a common value used by Bayesian modelers. It is completely arbitrary, just like the conventional 95% value used in frequentist approaches, with the exception that it the final prime number before 100 (McElreath, 2018).

graded inference. If the CI falls completely outside of the ROPE, evidence for a given effect is said to be robust (Kruschke, 2011). If the CI partially overlaps with ROPE such that one end falls within the ROPE and other falls outside it, then there is evidence to rule out parameter values on the other side of the ROPE (e.g., a CI partially within ROPE and partially outside the positive end; one would say there is evidence to suggest the effect is greater than or equal to zero, or that the sign of the effect is non-negative). If the ROPE entire contains the CI, then that is evidence in favor of accepting a null effect. Finally, if the CI spans the ROPE but also contains value outside both ends of it, then the evidence is labeled ‘equivocal’ (roughly equal evidence for positive and negative signs). These inferential scenarios are depicted in the Appendix (Figure A1.1).

For the current analysis, a ROPE of [-0.095, 0.095] was defined. These values reflect a 10% expected change in the likelihood of flipping over a card after transforming logistic regression coefficients back into the odds scale. This 10% threshold was selected because anything less than this value would translate to a change of 5 choices (i.e., moving from safe to risky) on the CCT for a hypothetical subject with 200 trials. Because most subjects did not play the game such that they experienced 200 trials, this means that anything less than 10% would result in a change of only 1, 2 or 3 risky decisions. This was not meaningful to me for this study.

Results

The following results are organized by effects of interest. I first report overall social decision-preferences (i.e., decision-making behaviors as a function of condition). I then report the moderating effects of relationship quality, reward type, and demographic variables (age, sex).

Overall Social Decision Preferences. I first examined results of the within-subjects only model (Table 1.1, Model 1) to understand population-level social decision preferences regarding parent and friend outcomes across all participants in the sample. As reported in Table 1.1 and shown in Figure 1.2, I observed robust evidence for a positive association between the Condition indicator variable (i.e., who gains and who loses on the current decision) and the likelihood of flipping a card.

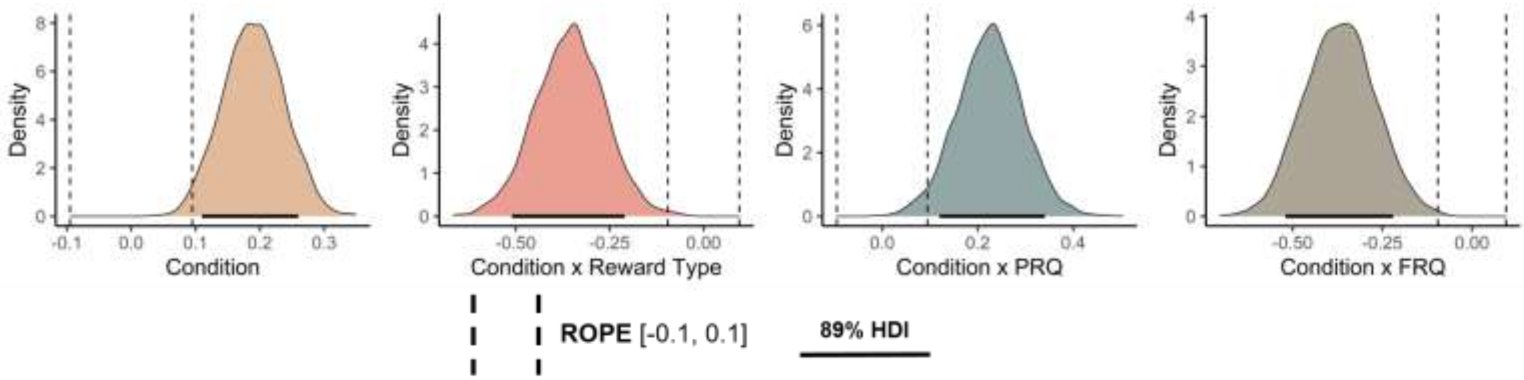
Table 1.1 Posterior estimates and credible intervals for hierarchical logistic regression models.

Term	Model 1	Model 2	Model 3	Model 4
Intercept	2.22 [2.13, 2.32]	2.22 [2.13, 2.31]	2.24 [2.04, 2.42]	2.19 [1.99, 2.39]
Condition	0.19 [0.12, 0.27]	0.19 [0.12, 0.26]	0.19 [0.12, 0.26]	0.33 [0.17, 0.47]
Return	0.07 [0.06, 0.07]	0.07 [0.06, 0.07]	0.07 [0.06, 0.07]	0.07 [0.06, 0.07]
Risk	-0.07 [-0.07, -0.06]	-0.07 [-0.07, -0.06]	-0.07 [-0.07, -0.06]	-0.07 [-0.07, -0.06]
PRQ	-	-0.22 [-0.36, -0.09]	-0.22 [-0.36, -0.09]	-0.22 [-0.35, -0.08]
FRQ	-	0.22 [0.02, 0.40]	0.22 [0.04, 0.43]	0.21 [0.04, 0.42]
Age	-	-	-0.02 [-0.07, 0.02]	-0.02 [-0.07, 0.02]
Sex	-	-	0.02 [-0.17, 0.23]	-0.01 [-0.21, 0.22]
Reward Type	-	-	-0.06 [-0.25, 0.15]	0.07 [-0.13, 0.26]
Condition x PRQ	-	0.22 [0.11, 0.34]	0.22 [0.11, 0.33]	0.22 [0.11, 0.33]
Condition x FRQ	-	-0.35 [-0.49, -0.18]	-0.35 [-0.51, -0.19]	-0.36 [-0.51, -0.20]
Condition x Age	-	-	-	-0.00 [-0.04, -0.03]
Condition x Sex	-	-	-	0.08 [-0.08, 0.26]
Condition x Reward Type	-	-	-	-0.38 [-0.54, -0.23]

Note. Condition was coded 0 = friend gain/parent lose, 1 = parent gain/friend lose; Return (expected value of a risky choice) ranged from -60 to 16.88 and was mean centered; Risk (standard deviation of outcomes associated with risky choice) ranged from 9.68 to 40.00 and was mean centered; PRQ and FRQ respectively represent parent and friend relationship quality; Sex was code 0 = male, 1 = female; age was mean centered; reward type was coded 0 = simulated, 1 = real. Values in brackets represent 89% HDI posterior distribution credible intervals.

The Condition parameter estimate is interpreted as a 20.92% increase in the expected odds of flipping over a card in the “parent gain—friend lose” scenario compared to the opposite (“friend gain—parent lose”). The parameter value has a parallel interpretation regarding safe decisions, such that participants are more likely to *pass* during the “friend gain—parent lose” condition. This is evidence for a population-level parent-over-friend social decision preference. Notably, the effects of Return and Risk are consistent with outcomes of other risk-return decompositions of the CCT (van Duijvenvoorde et al., 2015), which show individuals are more likely to flip over cards with increasing return and less likely to do so with increasing risk. It is important to consider that these parameter estimates are for 1 point's worth of return or risk, despite the fact that the HDI for these effects falls within ROPE. This is notable because risk and return values could change drastically over the course of several decisions within a given deck and therefore have compounding effects on decision behavior (e.g., return values could diminish nearly 90% during the course of a deck). In light of this, and because the sign of the effects were in the expected direction, I argue the effects of risk and return are an affirmative manipulation check that suggests participants were completing the task appropriately[8].

Figure 1.2. Posterior distributions for select model parameters



Note. 89% HDI refers to posterior credible intervals; posterior samples within the dashed lines (ROPE) were considered practically equivalent to zero; the histogram for the main effect of condition reflects posterior samples from Model 1, whereas samples in all other histograms are from Model 4; Condition x Reward Type refers to the effect of reward type on social decision preference; Condition x PRQ/FRQ refers to the effect of parent (P) or friend (F) relationship quality on social decision preference (positive values reflect a parent preference, negative values reflect a friend preference).

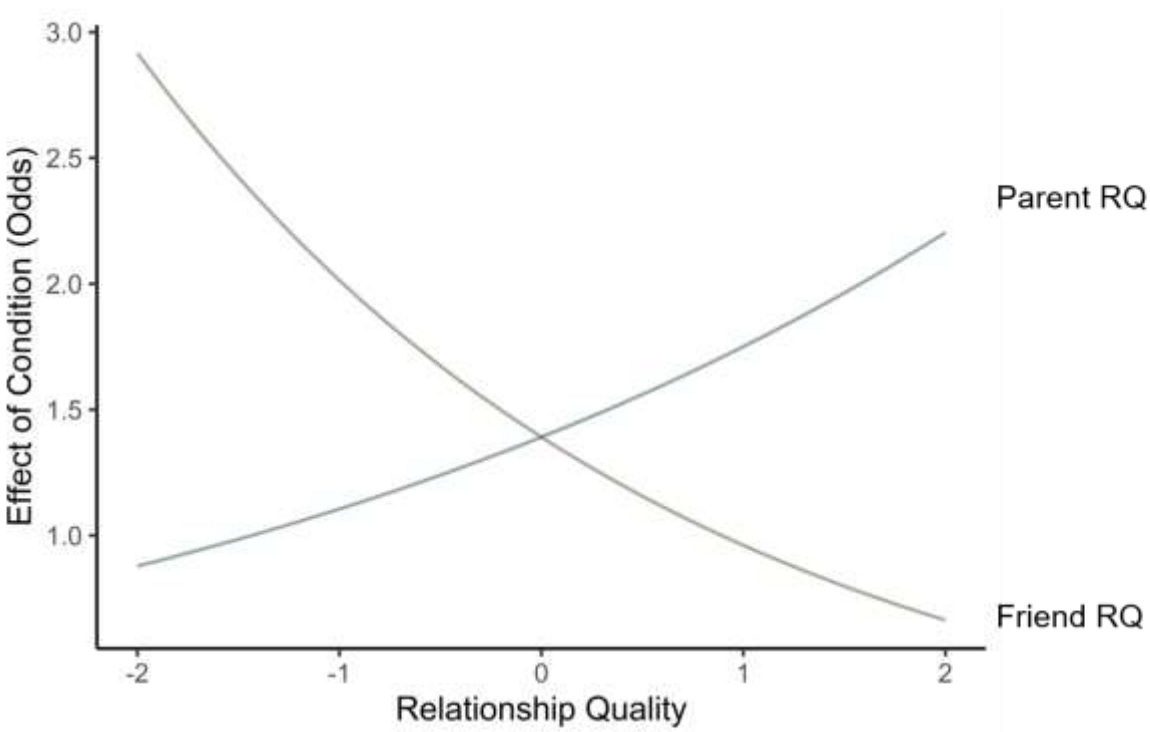
The Moderating Influence of Relationship Quality on Social Decision Preferences.

Despite a clear group-level preference, there was still considerable heterogeneity in individual

[8] A preliminary maximum likelihood analysis suggested that Context does not interact with the other level 1 variables (return, risk), suggesting that social decision preferences are invariant (or at least minimally sensitive) to different combinations of return and risk values.

decision preferences (see Figure A## for a visual depiction of group-level social decision preferences). I modeled this variability as a function of parent and friend relationship quality. For conceptual clarity, I interpret these moderators as influences on the effect of condition (as opposed to thinking about these terms as cross-level interactions). Also reported in Table 1.1 (Model 2), I observed a direct association between the effect of condition and parent relationship quality, and an inverse association between the effect of condition and friend relationship quality (Figure 1.2)

Figure 1.3. Unpacking the effect of relationship quality on social decision preferences



Note. RQ refers to relationship quality as measured by the IPPA Parent and Peer Attachment; Relationship quality was mean centered; Parameter estimates from Model 4 are used here in this figure. Parameter estimates were transformed back into the odds metric to facilitate interpretation of the outcome, binary decisions; the slopes therefore appear non-linear, given the multiplicative interpretation of odds, despite the coefficients being derived from a linear model. This plot assumes all other predictors in the model are set to zero. Notably, a *positive value* for the effect of Condition represents a *parent preference*, a *negative value* represents a *friend preference*, and *zero* represents *no preference* (i.e., equivocal).

Individuals with greater parent relationship quality were more likely to evince a relatively stronger decision preference towards their parent (statistically adjusting for friend relationship quality); individuals with greater friend relationship quality were more likely to endorse relatively stronger decision preferences favoring their friend (statistically adjusting for parent relationship quality). In absolute terms, however, Figure 1.3 shows that individuals with the highest friend relationship quality still showed equivocal social decision preferences. Parent and friend relationship quality were only modestly correlated (r = 0.229), indicating that statistical adjustment of each term in the model was appropriate and descriptive of actual participants in the sample (Miller & Chapman, 2001). It is therefore important to consider the effects in the context of parent-preference reference point, with parent and friend relationship quality serving to further entrench parent preferences or push individuals towards equivocal preferences. These results when adding additional terms to the model (Models 3, 4).

Reward Type, But Not Sex or Age, Affects Social Decision Preference. Interestingly, I observed robust evidence that indicated reward type moderated social decision preferences (Table 1.1, Model 4; Figure 1.2). Participants were more likely to favor their parent at the expense of a friend when rewards were simulated. By contrast, preferences were virtually equivalent when real rewards were at stake. I did not observe robust evidence to suggest that social decision preferences varied as a function of age or sex (Table 1.1, Model 4). The effect of age was essentially equivalent to the null value of zero, indicating stability of social decision preferences from late adolescent to young adulthood. There was some evidence to suggest female, relative to male participants may display a slightly larger parent-over-friend preference, but I can only definitively say the opposite is not true.

Interim Discussion 1

Study 1 aimed to help enhance the external validity of social decision-making research by examining whether decision preferences varied as a function of the social agents impacted by said decisions. I was specifically interested in understanding how individuals navigated decisions involving trade-offs for close others, as these types of choices tend to be the most consequential in everyday life. I found that a relatively large sample of individuals in the transition between late adolescence and young adulthood were more likely to favor a parent over a friend when making decisions that posed conflicting outcomes for each person. That a population-level preference emerged is significant for several reasons.

First, it is a simple proof of concept that shows that the identity of one's social decision targets influences decision preferences. Though straightforward, this finding can potentially have profound implications for nearly every area of research falling under the umbrella of social decision-making. They show that social decisions preferences are not monolithic and underscore the notion that social behaviors occur at varying levels of granularity. To the extent that our field wants its science to describe the real-world phenomena it studies, a parallel effort must be made to understand this granularity. This is not to say extent literature need to be discarded, but rather that more focus ought to be placed on obtaining a specific understanding of social decision-behavior in varied contexts. Whereas this could generally be said for other social cognitive phenomena, social decision-making is arguably more sensitive to inter-situation differences (e.g., whom is affected by one's decisions) because decision behavior is not necessarily as generalizable as more basic social cognitive phenomena. In other words, decision-making is arguably more sensitive to situation-level idiosyncrasies than a phenomena such as theory of mind or social rejection[9].

Second, this finding opens to door to further questions. Why do individuals favor parents over friends? What features of these relationships or social agents beget decision preferences? Relationship quality findings offer an important clue, but relationship quality itself is a complex construct that emerges from, and is correlated with, a variety of psychological processes (Armsden & Greenberg, 1987). This means a statement such as ‘relationship quality drives social decision preferences’ is ultimately incomplete—it merely indicates that socioemotional motivations affect social decision preferences without any satisfying information about mechanistic specificity. Studies 2 and 3 will attempt to address this issue.

Third, these findings may be of interest to the developmental psychology community. Bucking popular lay theories of adolescence and young adulthood (Arnett, 2000; Steinberg & Morris, 2001), the findings from this study suggest parent relationships are so important that young adults are willing to incur losses for their friends to benefit their parents. This expands existing work in developmental psychology that shows such individuals identify with their parents and hope to meaningfully contribute in the context of family relationships (Fuligni, 2018; Tsai et al., 2013). This could be because parents may be the most stable relationships at a time of inherent relationship instability (Arnett, 2014), or because individuals want to give back to their caregivers after a lifetime of emotional and financial support (Fuligni, 2018; Fuligni & Pedersen, 2002). As an aside, contextualizing these findings in the broader developmental literature is important because it provides additional information about the *why* behind social decision

[9] This isn’t to say inter-situational differences don’t exist with these phenomena—they most certainly do. The point I argue is that social decision-making is *more* sensitive to such differences.

preferences. If social decision studies begin to investigate the origins of increasingly granular decision preferences, contextual work from adjacent fields like that discussed here could prove to be invaluable.

The interesting moderation findings merit also merit discussion. With respect to age, it was somewhat surprising to see there were no age effects in our sample. Although the age range was somewhat truncated, one could have still argued that older individuals would have displayed stronger parent-over-friend preferences given developmental studies that show parent-child cohesion improves as individuals age out of adolescence (Tsai et al., 2013). Instead, our study provides evidence that these preferences are developmentally stable within the range of late adolescence and early adulthood. Our findings were similarly curious with respect to reward type, with real rewards being associated with only slight parent-over-friend preferences (whereas simulated rewards were associated with clear parent-over-friend preferences). It could be that simulated rewards promote the pursuit of idealized goal structures, or the relative prioritization of various goals in absence of salient external demands. Real rewards force individuals to modulate their goal structures according to reality's demands (Freitas et al., 2004), suggesting participants in the simulated condition were making decisions according to an ideal goal structure in which parents are more important than friends. However, once rewards became real, participants were forced to face the reality that losses were no longer immaterial and may have had difficulty making decisions that rigidly conformed to an ideal goal structure. Alternately, it was also possible that earning real rewards elicited greater emotional salience, blunting the cognitive skills needed to pursue these idealized goals (Freitas et al., 2004).

Study 2: Neural Representations and ties to Social Decision-making

Study 1 served both as a proof of concept—showing social decision preferences vary as a function of the implicated social agents—as well as a substantive finding that late adolescents in the target population appear to prioritize parents over friends during social decision-making. Study 2 builds off study by examining whether value-based expression of neural representations relate to social decision-making involving close others.

That social decision-preferences varied as a function of relationship quality with the affected social agents suggests that individual socioemotional motivations influence social decision behavior at the individual level. These findings beg the question of what precisely drives these preferences. Because preferences changed as a function of the implicated social partner, and not decision-level features (e.g., return, risk, etc.), this implies that features of these relationships drive preferences. This suggestion is notable because most prior work in social decision-making research examines how individuals compute latent quantities (e.g., value) over decision-level inputs; by comparison, studies have spent considerably less effort understanding how agent-level inputs are computed or processed. Study 2 builds upon the previous study by using fMRI to measure structural features of parent and friend representations—specifically focusing on how representations are expressed as signatures of value—and link these features back to social decision tendencies.

For over a decade, social decision-making work has leveraged fMRI to glean insights into the nature of decision processes (Rilling & Sanfey, 2011). Cognitive neuroscience can help answer mechanistic questions that models of decision-making cannot quite articulate with purely psychological means (Chang et al., 2011; Harris et al., 2018; Sokol-Hessner et al., 2012, 2013;

Sokol-Hessner & Rutledge, 2019). For instance, prospect theory[10] postulates that individuals tend to overweight losses relative to gains (Tversky & Kahneman, 1992). It remained unclear for decades whether loss aversion was driven by top-down cognitions or bottom-up affective responses. It was not until scientists used fMRI that they were able to determine that loss averse attitudes are largely driven by bottom-up affective processes in the amygdala (albeit with top-down modulation from lateral prefrontal cortical regions) (Sokol-Hessner et al., 2012, 2013). Importantly, this example also serves to illustrate how most imaging work in this arena has attempted to uncover the neural underpinnings of cognitive computations performed over decision-level features (e.g., appraisal of subjective value, ambiguity, etc.). This and other prior work have revealed that social decisions are made according to value-based rules – decision alternatives are selected when they maximize *subjective* value.

Very little work has examined how the brain represents agents implicated in social decision-making and whether said representations are consequential for shaping the value-based calculations that frequently guide decision preferences (Fareri et al., 2020). This is somewhat surprising, given that separate research arcs in social neuroscience that have accumulated a rich literature on neural representations of social agents (Guthrie et al., 2021; Hassabis et al., 2014; Huth et al., 2016; Parkinson et al., 2017; Wang et al., 2017). Basic cognitive neuroscience research shows that neural representations of objects and places affect interactive behavior with the physical world (Charest et al., 2014). Moreover, neural representations form the bases for psychological representations (Amodio, 2019; Huth et al., 2012, 2016), which are in turn crucial for guiding behavior (Tamir & Thornton, 2018; Tversky & Hutchinson, 1986). Additional work

[10] Prospect theory was initially conceptualized as a theory for financial decision-making, but has since been applied in social domains (see Sokol-Hessner & Rutledge, 2019).

from social neuroscience has also shown that salient social information about other agents (e.g., social network position, reciprocity behavior) is encoded via activity reward-processing systems such as the ventral striatum and medial prefrontal cortex (Braams & Crone, 2016; Delgado et al., 2016; Guthrie et al., 2021; Hackel et al., 2015; Morelli et al., 2018; Zerubavel et al., 2015). Together, these two lines of research support the notions (i) that representations of social agents figure to be critical for shaping decision preferences involving said agents and that (ii) value-based processes are likely integral components of social representations.

Current Study. Whereas Study 1 demonstrated that individuals tend to favor parents over friends on average, it did not reveal why. Study 2 used fMRI to determine the extent to which neural representations of close others (parents and friends) were encoded as neural signatures of valuation, and related these estimates to social decision preferences. Although traditional social neuroscience approaches to examining neural representations are based on semantic knowledge (e.g., measuring representations when semantic information related to social targets is activated) (Chavez et al., 2017; Guthrie et al., 2021; Hassabis et al., 2014; Wang et al., 2017), I explicitly focused on a value signature here for two reasons. First, as previously mentioned, strong evidence suggests that human decision-making follows a value-based architecture (Sokol-Hessner & Rutledge, 2019), and emergent work suggests that social representations of others does indeed depend on value-based neural circuitry (Delgado et al., 2016). This architecture is so important as to be observed in non-human species (Chen et al., 2006), underscoring the need to understand the role of neural representations from a value-based perspective. Second, and relatedly, the construct of value is emergent and is supported by multiple constituent psychological processes (Boer & Boehnke, 2015; Higgins, 2015; Niv & Chan, 2011; Zhou et al., 2019). The notion of value is not purely synonymous with maximizing an external reward.

Instead, psychological theories of value posit that valuation is a process that requires appraisal of internals states (desires, needs) against realities of the external environment (Boer & Boehnke, 2015; Higgins, 2015). Psychologically, for something to be valued, it must quench one's basic needs and desires. This is especially true in the context of close relationships, as work shows that core features of close relationships (e.g., attachment), appear to also be encoded in value-based brain regions (Tottenham, 2020). In this way, valuation is reflective of many different psychological processes, underscoring the idea that value-based encoding of social relationships reflects intrinsic evaluation of a given relationship partner's ability to sate internal needs and desires.

Because social decision-making appears to follow a value-based architecture (Guassi Moreira et al., 2020), it follows that social agents whose neural representations closely follow a neural signature of valuation are more likely to be prioritized during decision-making. This is not to say that individuals may cherish one close other more over another—rather one close other may help satisfy dominant internal goals better than another. The representation of close others may contain information that comprises the bases of such goals and therefore spontaneously computes value signals. In other words, value-based computations are an integral part of one's representation of close others because value is synonymous with the very features and qualities of a close relationship that enable closeness in the first place.

Hypotheses. The primary hypotheses for this study were that individual neural representations of parents—relative to friends—will be more strongly expressed as neural signatures of value, and that individual differences in pattern expression values will track with social decision preferences (i.e., greater value expression in one's neural representation of a close other will be associated with a greater tendency to favor said close other).

Method

Participants. Participants for this study were comprised of 48 late adolescents (18-19 years) from the West Los Angeles area in the United States. Considerations regarding sample size justification are enumerated below in a separate section. Recruitment practices involved posting flyers around the UCLA campus and sending mass emails to the freshmen and sophomores in the study body. In order to be eligible to participate, individuals were required to (i) be between the ages of 18 and 19 years old (i.e., late adolescents), (ii) be eligible for MRI scanning (e.g., no metal implants, no claustrophobia, etc.), (iii) be a fluent English speaker, (iv) have no neurological impairments, (v) be able to nominate two close others (a parent and friend) and provide custom stimuli of each (photographs, names). Participants were compensated with $25 (USD) cash payment plus an additional $1-5 bonus (amount chosen randomly, described in greater detail below). Three participants were excluded outright from all analyses (one because of a scanner computer error, a second due to poor overall data quality, and a third due to the discovery of a biological artifact), rendering the final sample size equal to 45. All participants provided written consent in accordance with the policies of the UCLA Institutional Review Board.

Sample Size Considerations. The best practices for determining sample size for Study 2's proposed analyses were unclear due to poorly defined procedures for estimating statistical power in fMRI research (G. Chen et al., 2017; Poldrack et al., 2017). To further complicate matters, fMRI is a particularly expensive neuroimaging modality, which constrained my flexibility for collecting many data. In light of these realities, I first addressed the issue by identifying the median cell in size in human neuroimaging using fMRI. The most recent report on this topic concluded the median size was N = 35 (Poldrack et al., 2017). After evaluating the availability of funding, I realized I had the means to scan at least 35 participants. I then raised my target *N* until I exhausted all funding that could be flexibly used for a project of this caliber, corresponding to *N* = 50. The logic for increasing the sample size was that larger sample sizes are generally better for statistical analysis and the median cell size in this era of neuroimaging does not necessarily reflect adequate statistical power any given type of analysis. Put differently, I set out to scan as many participants as possible barring financial constraints and those eventually imposed by the COVID-19 pandemic. This resulted in 48 participants, 45 with usable data.

Experimental Protocol.

Overview. Here I outline the general experimental protocol that was performed. Upon recruitment, participants were asked to nominate a parent and close friend of their choice, and provide stimuli (photos, names) of each person prior to their scheduled scan date. Participants underwent scanning at UCLA's Ahmanson-Lovelace Brain Mapping Center. Forty-five minutes prior to scanning, participants were trained to complete computer tasks for the scanner and verified a final safety screening. Afterwards, participants underwent an fMRI session where they completed a series of computer tasks while being scanned. Last, participants completed a post-scan session comprised of administration of additional computerized tasks and a self-report survey. I now describe each element of this procedure in greater detail.

Parent-Friend Nomination and Stimuli Collection. Upon signing up for the study, participants were informed that the study involved making hypothetical decisions on the behalf of a parent and close friend, and that they must nominate one of each. Additional constraints were imposed for friend nomination: participants were not allowed to nominate friends who were current romantic partners or family members in effort to avoid potential confounds. Afterwards, participants were told they must provide specific stimuli of each parent and friend. Five 'passport style' headshots from five different camera angles of each close other, in addition to the name the participant used to address them, were required prior to completing the scan. Images required neutral facial expressions, both eyes to be open, mouth shut, eyes locked straight ahead, and no head tilt. The experimenter first reviewed these requirements with participants via telephone and then sent them a PDF file with complete, detailed instructions. The experimenter assessed images for quality prior to the scan and asked participants for re-shoots if images did not comply with requirements.

Figure 2.1. Sample example stimuli.

Note. These images were taken from the sample PDF instruction file. These images are not actual subject data.

fMRI Tasks.

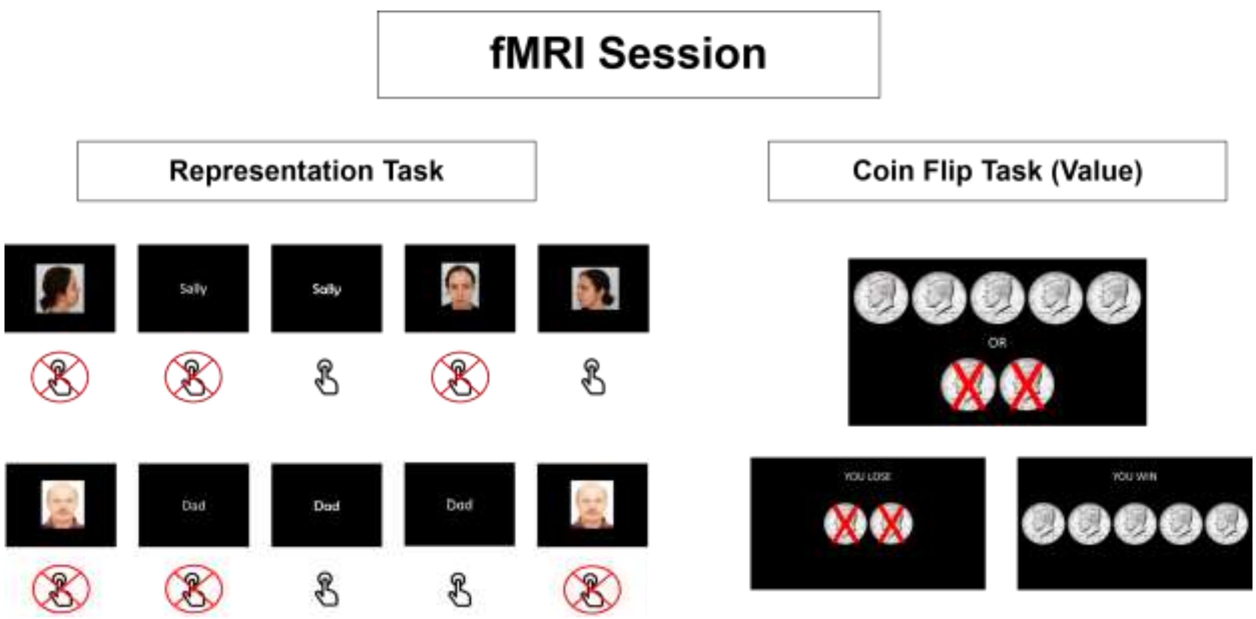
Parent-Friend Representation Task. In order to elicit spontaneous neural representations of parents and friends, participants completed a Parent-Friend Representation Task. Participants were shown stimuli related to their parent and friend in a block design in order to capture neural representations of each close other. In a given block, participants saw random ordering of stimuli pertaining to one close other. These stimuli were comprised of the aforementioned five headshots in addition to the close other's name[11] printed in five unique fonts – 'Berlin', 'Broadway',

[11] Based on feedback received during pilot scans, I asked participants to provide the labels they use to address each close other (e.g., 'Mom' or 'Dad' for a parent).

‘Calibri’, ‘Colonna’, and ‘Comic Sans’ (10 unique stimuli). Each block contained 20 rapid presentations (2 for each stimulus) of said stimuli (1000ms) with a brief ISI between images (500ms). Participants completed a one-back based on stimulus type (photo vs text) to ensure they were paying attention (i.e., press a button if the current stimulus type, photo or text, matches the one shown just before it). 15000 ms of fixation between blocks was presented to account for lagged effects of the hemodynamic response function. Six blocks (3 parent, 3 friend) and six interblock fixation periods were presented per run. As a result, the entire task lasted approximately 4.5 minutes (270s): [1500ms/trial x 20 trials/block x 6 blocks] + [15000ms interblock fixation periods x 6 fixation periods]. Crucially, the use of varying photographic and text stimuli theoretically allows one to elicit *amodal,* spontaneous neural representations of parents and friends. This helps safeguard against the risk of basic perceptual confounds. Various elements of this task were designed to be broadly consistent with prior literature (Gee et al., 2014; Parkinson et al., 2017; Taylor et al., 2009; Zerubavel et al., 2015).

Coin Flip Task. Following the representation task, participants completed two runs of a commonly used reward-task (Braams & Crone, 2016). During this event-related task, participants guessed about the outcome (‘Heads’ or ‘Tails’) of a series of coin flip gambles in order to win or lose monetary rewards (presented as coins). Each trial began with a reward summary (3000ms), a screen that details the amount awarded or lost for guessing correctly or incorrectly, respectively. Participants were required to make their guess, via button press, at this stage (‘Heads’ or ‘Tails’). Following a 1000ms ISI, participants received feedback about whether their guess was correct or incorrect (2500ms). A jittered ITI separated trials, with values drawn from an exponential distribution (mean = 2880ms, SD = 2660ms, range = 1000-10000ms). Each run lasted approximately 6 minutes. Participants completed 30 trials per run, broken down across three distinct trial types: (i) win 3 coins, lose 3 coins; (ii) win 5 coins, lose 2 coins; (iii) win 2 coins, lose 5 coins. Participants were told the coin is fair (i.e., P('Heads') = ½), but in reality, the game will be rigged such that individuals won and lost approximately half of their guesses. This is to ensure there are enough gain and loss events for subsequent modeling and estimation purposes. To obtain a relatively generalized signature of valuation, one run varied the type of coins (Kennedy coin vs Sacagawea coin) and thus the perceptual features of the coin (color: silver vs gold; gender of the head: male vs female; etc.). The orientation of the coin also varied for this reason (i.e., half of the reward summaries showed the coins on the 'Heads' side, the other half showed them on the 'Tails' side). Last, participants were informed a subset of the trials would be selected at random and added to, or subtracted from, their earnings (up to +/- $5). In actuality, participants always received randomly selected bonus between $1 - $5.

Figure 2.2. Schematic of the two fMRI tasks.



Note. The four runs of the representation task were always administered before the two runs of the coin flip task.

fMRI Data Acquisition. Neuroimaging data were collected using a research-dedicated 3 Tesla, Siemens Magnetom Prisma MRI scanner and 32-channel head coil. A high resolution T1* magnetization-prepared rapid-acquisition gradient echo structural image was acquired for registration purposes (MPRAGE; TR = 2400ms, TE = 2.22ms, Flip Angle = 8º, FOV = 256 mm^2, 0.8 mm^3 isotropic voxels, 208 slices, A >> P phase encoding). Functional runs were comprised of T2*-weighted multiband echoplanar images (TR = 1000ms, TE = 37ms, Flip Angle = 60º, FOV = 208 mm^2, 2.0 mm^3 isotropic voxels, 60 slices, A >> P phase encoding, multi-band acceleration factor = 6). These parameters were obtained by surveying recently published studies that employed similar analytic techniques (Chang et al., 2015; Chavez et al., 2017).

Post-Scan Measures.

Participants completed the following post-scan measurements following the scan.

Modified CCT. As in Study 1, I administered the modified CCT to participants to assess social decision-making preferences. The task parameters, administration, and instructions were exactly the same as Study 1. Only hypothetical rewards were used.

Relationship Assessments. I assessed various aspects of parent-friend relationships in a self-administered survey. First, participants completed the same salience procedure as described in Study 1, writing about memories they have with their parent and friend, as well as describing them using a handful of words and phrases. For this study, the procedure was not intended to amplify the salience of making hypothetical decisions for a parent and friend and instead it was administered to collect more data about these two close relationships. Afterwards, participants completed the IPPA, as described in Study 1. Participants then indicated how often they spent time with their nominated parent and friend in a typical month ("During an average month, how often do you have the chance to spend time with the [parent/friend] you nominated?") along a 7-point Likert scale (1 = "0 Days', 2 = "1 or 2 Days", 3 = "3 to 5 Days", 4 = "6 to 9 Days", 5 = "10 to 19 Days", 6 = "20 to 29 Days", 7 = "All Days").

Additional Measures. Participants completed a set of additional measures collected with the intent of answering current and future exploratory research questions. These measures included two computerized assessments: the Social Gambling Task (Kwak et al., 2014), used to measure parent and friend preferences in the probabilistic learning domain, and the Cups Task (Levin & Hart, 2003), a monetary self-oriented risk-taking task that captures individual sensitivities to reward and uncertainty (Uy & Galván, 2017). The measures also included other self-report assessments: The domain specific risk taking scale "DOSPERT" (Figner & Weber, 2011), the brief sensation seeking scale (Hoyle et al., 2002), and the substance use items from the US Center for Disease Control and Prevention's Youth Risk Behavior Survey (consistent with prior work in this age group (Telzer et al., 2014).

Analysis Plan

Overview. In order to determine how value-based neural representations relate to social decision-making involving close others, I conducted a pattern expression analysis with the fMRI data (Doré et al., 2017; Hong et al., 2019). Pattern expression analyses are used to answer questions about how strongly a given brain state is expressed as a psychological process of interest. For this study, the intent was to determine how strongly neural representations of parents and friends were expressed as signatures of value. I first tested whether parent representations are more strongly encoded as value signatures than friend representations. Afterwards, I related individual differences in pattern expression scores with social decision-making preferences on the modified CCT.

fMRI Data Preprocessing. Prior to preprocessing, data were visually inspected for artifacts and anatomical abnormalities. I preprocessed and analyzed the data using the fMRI Expert Analysis Tool (FEAT, Version 6.00) of the MFRIB Software Library package (FSL, Version 5.0.9; fsl.fmrib.ox.ac.uk). Preprocessing began by using the brain extraction tool (BET) to remove nonbrain tissue from functional and structural images, followed by head motion correction via spatial realignment of functional volumes using MCFLIRT. The data were hi-pass filtered to remove low frequency artifacts (45s for the parent friend localizer; 100s for the coin flip task). From there, the extent of head motion artifacts was estimated by using the FSL Motion Outliers command to document volumes that exceed a 0.9 mm threshold of framewise displacement (FD; Siegel et al., 2014). Runs with 25% of volumes exceeding this threshold were excluded from analysis. To help reduce high frequency noise introduced by realignment (Etzel et al., 2011; Misaki et al., 2014), data were smoothed with a 1 mm Gaussian kernel (full width at half maximum). Data were pre-whitened prior to analysis to correct autocorrelated residuals. FSL's boundary based registration algorithm (Greve & Fischl, 2009) was used to register functional data to the high resolution structural scan (MPRAGE). MPRAGE images were then nonlinearly registered to the MNI152 template image (10-mm warp resolution), and the ensuing transformation matrix was used to register functional images to standard space. This step also resampled voxel size to 2mm^3 isotropic.

All participants had usable data for the parent and friend representation eliciting task, although three participants only had 1, 2 and 3 usable runs (out of four), respectively, of the task available for analysis. Three participants were excluded from analyses involving the coin flip task. Two such participants were excluded because they lowered part of their heads out of the coil during the coin flip task, rendering missing data for large parts of the temporal pole. The third such participant was excluded due to head motion, as they averaged 22 volumes exceeding the FD threshold (average maximum FD = 8.82mm) across both runs.

Overall, head motion in this sample was remarkably low. Averages of *n* volumes exceeding the FD threshold and maximum FD values were computed within each subject for each fMRI task. The means of these intra-subject averages was used as a descriptive metric to reflect the overall head motion in the sample. The mean intra-subject average of *n* volumes exceeding the FD threshold on the parent-friend representation eliciting task was 0.484mm. The mean intra-subject average of the maximum FD value for this task was 0.620mm. Substantively, this means an 'average subject' is expected to move less than one volume above the FD threshold per run, and that their maximum FD value per run is expected to be ~ 0.6mm. Only fifteen subjects exceeded the FD threshold during any run of the parent-friend representation eliciting task. For the coin flip task, the mean intra-subject average of *n* volumes exceeding the FD threshold was 0.467, and the mean intra-subject average of the maximum FD value was 0.637mm[12]. Twelve subjects exceeded the frame displacement threshold during any run of this task, whereas the rest did not.

Multivariate Pattern Estimation. I estimated three multivariate neural patterns: a parent representation, a friend representation, and a value-based signature.

Estimating the parent and friend neural representations was accomplished by modeling the parent and friend representation eliciting task with a standard General Linear Model (GLM) analysis. Each run of the task was submitted to a fixed effects GLM analysis in FSL. Parent and friend blocks were modeled with respective boxcar regressors, convolved with the hemodynamic

[12] This estimate excludes the aforementioned outlying participant who averaged 20+ volumes exceeding the FD threshold.

response function (double gamma) and bandpass filtered to avoid reintroducing noise into the data. Slice timing effects were addressed by also modeling the temporal derivative of each task regressor. Head motion was statistically adjusted for by adding rotation and translation parameters, along with their derivatives and squares (obtained from MCFLIRT motion correction) as nuisance regressors. To further adjust for potential spurious effects of head motion, I included additional regressors for individual volumes that exceeded the 0.9 mm FD threshold. Two linear contrasts were computed: parent > baseline and friend > baseline. A second level analysis was carried out to average contrast estimates over the four runs, using a fixed effects model and forcing random effects variance to zero. The ensuing parent > baseline and friend > baseline maps, one each per subject, served as the estimates of parent and friend representations.

Creating a neural signature of value necessitated the use of machine learning methods employed by other groups tackling a similar task (Chang et al., 2015; Cosme et al., 2019; Wager et al., 2013; Reddan et al., 2018). Broadly described, this process involved training a statistical model to predict gain and loss values on each trial of the coinflip task based on brain activity. This process yielded a statistical map containing voxel weights that represent the strength of association between voxel activity and reward/loss outcomes. The first step in this task was to compute brain activity for individual trials on the coin flip task. I accomplished this by conducting a least squares single (LSS) analysis (Mumford et al., 2012, 2014). Briefly, a LSS entails creating a unique fixed effect GLM for every trial, in every run, across all subjects[13]. Every trial is modeled as a single-event regressor in its respective GLM, and all other trials are

[13] All other GLM specifications (e.g., slice timing correction via temporal derivatives, regressor convolution, etc.) for the LSS analysis were identical to those used in the parent-friend representation GLMs.

modeled as they traditionally would otherwise. For the coinflip task, this meant that any given LSS GLM would contain a regressor for the current 'target trial', a regressor for gain outcomes, a regressor for loss outcomes, and a regressor for guessing between 'Heads' or 'Tails' (i.e., the length of presentation time for the reward summary). A linear contrast comparing trial > baseline was estimated for each GLM. The ensuing single-trial estimates from all subjects were concatenated and used to extract a t x v matrix, containing brain activity during the t-th trial in the v-th voxel (whole brain). Given the high dimensionality of this matrix (209,036 voxels), principal components analysis (PCA) was employed to reduce the number of features (i.e., voxels). Finally, penalized regression (e.g., LASSO, ridge) models were fit to the data, predicting the monetary outcome of each trial from its brain activity and thus yielding a set of weights for each principle component. Weights for each component were backtransformed into voxel space, yielding the final neural signature of value. To ensure the signature was specific to value and did not inadvertently tap another psychological process, I cross-referenced its similarity with publicly available meta-analytic maps of similar and distinct constructs. More details about each of these individual steps is presented in the results section.

A second method was used to derive a neural signature of interest (rationale explained below). I used meta-analytic maps from the online Neurosynth platform (Yarkoni et al., 2011). Neurosynth is an automated tool that extracts coordinates of brain activity from an actively maintained database of 14,371 studies (last updated July 2018), extracts high frequency terms occurring in the database's studies, and uses this information to conducted a meta-analysis of said activations for each term. Two images are computed for each term: a uniformity image and association image. The uniformity map captures the degree of activity in the brain for a given term (comparable to how one would interpret results from a 'standard' whole-brain, univariate analysis). The association map is more selective, as it controls for base rates (e.g., quantifies how much more likely a given brain region is likely to be activated for a given term relative to studies that don't include that term). More detailed on the platform can be accessed at neurosynth.org/faq. For this study, I used the meta-analytic map for the term 'reward'. According to Neurosynth, 922 individual studies contributed to this term's meta-analysis when I downloaded its maps in 2020. Both maps, uniformity and association, were used in analyses reported below.

I used this two-pronged approach to capturing neural signatures because no one method was decisively better than the other was. The benefit of using a sample-specific signature is that it may be better calibrated to individuals in the current sample, ensuring that individual differences in valuation owing to the sampled population are appropriately represented in the signature. Indeed, recent work has documented the high degree of individual differences in neural signals and has advocated for the use of personalized imaging applications. On the other hand, using a meta-analytically defined image does not offer the potential benefits of personalization, but offers a different kind of precision in that it aggregates data based on thousands of participants.

Pattern Expression Analysis. Pattern expression analysis is an elegant way of capturing how much a given psychological process (neural signature) contributes to a representation or state. The analysis is straightforward: it simply involves taking the voxel-wise dot product between values in one's neural representation and neural signature of interest. The computation is given by the following equation.

$$\sum_{i=1}^{n} w_i x_i \tag{7}$$

Where n is the number of voxels, w_i are the weights of the neural signature (valuation, in this case), and x_i is the neural activity (measured via BOLD) from the representation's voxels.

Statistical Analysis. After extracting pattern expression scores, I first examined whether parent representations were more strongly encoded as signatures of value, relative to friend representations. I achieved by analyzing paired differences in parent – friend pattern expression scores in a standard Bayesian framework.

$$\text{Diff}_i \sim N(\delta^*\sigma, \sigma^2) \tag{8}$$

Diff_i, representing the paired pattern expression difference score for the i-th participant, was modeled as being drawn from a normal distribution, centered around a mean ($\delta^*\sigma$) and variance (σ^2). The mean was parameterized as $\delta^*\sigma$ so that resulting summary statistics reflect standardized effect sizes (i.e., the mean is expressed in terms of standard deviations). The variance was given a Jeffreys prior ($p(\sigma^2) \propto 1/\sigma^2$), and δ—the mean effect size—was modeled as being distributed Cauchy ($\delta \sim \text{Cauchy}(0, r)$, where $r = 1/\text{sqrt}(2)$). This analysis tested the first hypothesis, which addressed mean level differences in value-based pattern expression in neural representations of parents and friends.

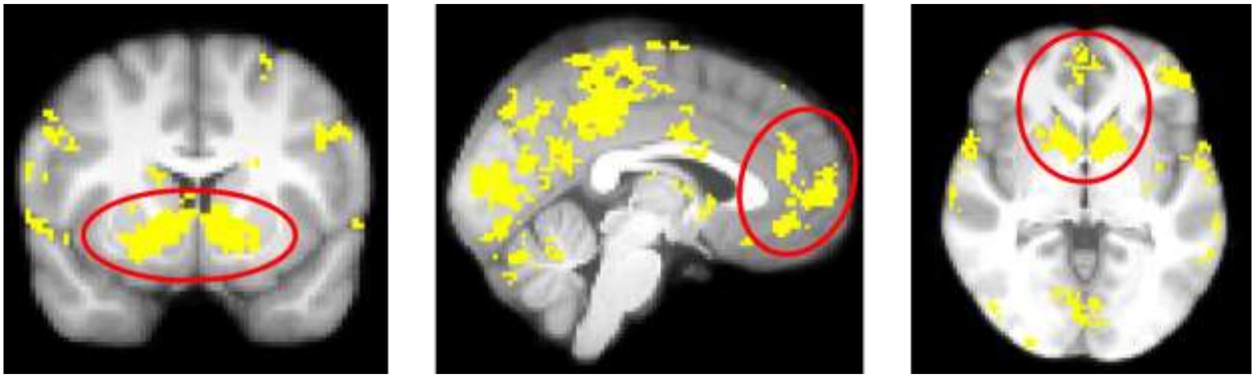
The next analytic step tested whether individual differences in pattern expression scores predicted social decision preferences using the same hierarchical Bayesian logistic regression framework. The form of these models in Study 2 is identical to Study 1, with the following exceptions. First, age and sex are not included in the model because (i) the age range is too narrow in this study (18-19 years) and (ii) there were no notable sex differences in Study 1. Second, parent and friend pattern expression scores take the place of parent and friend relationship quality scores in the between-person component of the model. Full equations for this model are omitted, given the high degree of similar between those from Study 1.

For all analyses here, the inferential criteria were the same as in Study 1. For paired analyses, the ROPE was expanded slightly to [-.1 to .1], as effects less than 1/10th of one standard deviation were of no interest.

Results

Manipulation Checks. Before beginning the analysis plan outlined above, I conducted two key manipulation checks. First, I computed linear contrasts (win > loss) from a traditional univariate analysis of the coin flip task to ensure the task was recruiting brain regions previously implicated in reward processing (Haber & Knutson, 2009; Knutson et al., 2001). Cluster-corrected (Family-wise-error < .05, cluster defining threshold Z > 3.1) results using random field theory show robust activation in the ventral striatum (k = 1570, L: x = -14 y = 6 z = -10, Z = 5.516 ; R: x = 16 y = 4 z = -12, Z = 5.517) as well as the medial prefrontal cortex (k = 539, x = -4 y = 56 z = 2, Z = 4.500) for the Win > Loss contrast, indicating a successful replication of prior work (Haber & Knutson, 2009; Knutson et al., 2001). Figure 2.3 visualizes these results. Importantly, this result indicates that I can proceed with using the coin flip task to build a neural signature of valuation since it does it appears that the correct psychological process was evoked.

Figure 2.3 Results of Win>Loss contrast during the coin flip task (reward).



Note. Winning, relative to losing, on the coin flip task evoked robust activity in the ventral striatum and medial prefrontal cortex (circled in red). Cluster corrected (Family-Wise-Error of $p < .05$) using FLS's FLAME1 (Cluster Defining Threshold of $Z > 3.1$).

Second, I analyze the modified CCT data without any between person predictors, to check whether I was able to replicate the overall parent-over-friend preference. Using the same modeling framework as previously described, I indeed observed evidence for a mean-level parent-over-friend social decision preference (posterior mean of social decision preference parameter: 0.30, 89% CI = [0.16, 0.44]). This analysis was similarly critical because it suggests that any potential null effects in other analyses would not be due to the current sample exhibiting differing social decision preferences than that of Study 1.

Capturing a Sample Specific Neural Signature of Value. I began the process of creating a sample specific neural signature of value by concatenating single trial activations from all subjects in a *t* by *v* matrix, where each row represented a single trial from the *i*-th subject, and each column represented a specific voxel. The matrix was reduced using PCA, following the precedent established by other similar studies (Chang et al., 2015; Krishnan et al., 2016; Wager et al., 2013). I kept 1500 components, comprising 90% of the variance explained in the original matrix. LASSO (Least Absolute Shrinkage and Selection Operator) and ridge regression were used to predict each trial's monetary value from BOLD activity, indexed by the set of principle components. The difference between ridge and LASSO regression lies in the penalization term. The ridge regression penalty nudges parameter terms *towards* zero (l2 regularization), whereas the LASSO regression penalty can set parameters *to* zero (l1 regularization).

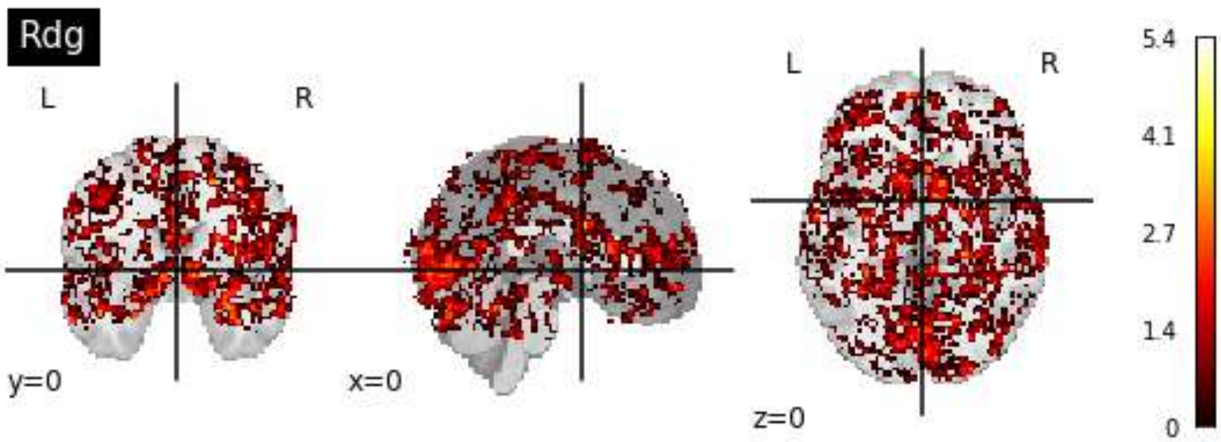
Ridge and LASSO regression were selected for several reasons. Foremost, the nature of the data demanded an analytic method that could handle continuous outcomes. Second, both methods use penalized estimators, which have the effect of regularizing parameter estimates (i.e., biasing them towards zero in order to reduce variability in sample-to-sample estimates). As previously discussed in this dissertation, this has the effect of enhancing generalizability of parameter estimates. In this case, this helps ensure that the sample-specific neural signature of value is not overfit to individuals in our sample (i.e., it is likely more reflective of the population it was derived from). Third, they are broadly consistent with existing, similar studies (Chang et al., 2015; Krishnan et al., 2016; Wager et al., 2013). Last, these models can handle highly parameterized models without encountering estimation problems (e.g., parameter estimate instability).

Both ridge and LASSO regression require a user-specified tuning parameter necessary for computing the penalty term (i.e., there is no analytic solution to determine the 'best-fitting' parameter). I used 10 fold cross-validation to determine the best penalty term to use for both ridge and LASSO models. Briefly, *k*-fold cross-validation is a technique from machine learning and statistics used for model selection and tuning. Data are partitioned into *k* folds and each fold is iteratively used as 'test' data to assess the fit of model parameters estimated that were 'trained' from the remaining data. The end of this process results in an average assessment of how a given model and its parameters will perform when predicting novel, unseen data. Operating under the assumption that a neural signature of any psychological process need be generalizable, cross-validation is thus an appropriate method for selecting the tuning parameter. After obtaining the ideal tuning parameter, I used both ridge and LASSO on the complete dataset to fit a model predicting monetary value on each trial from principle components of brain activity (indexed via the BOLD signal). Evaluating both models using the R^2 metric of model fit, I found that the ridge regression model fit the data better than LASSO (R^2 stats here), although both were similar.

I backtransformed the weights of each principal component into the original voxel space, and thresholded the weights at zero[14], creating the final neural signature map. I initially completed this procedure for both the LASSO and ridge models to visually the compare the ensuing models. Notably, the LASSO map was much sparser than the ridge map, in a manner that would have likely rendered it inappropriate for use in the current sample (because very few contiguous regions were formed by the model weights for the LASSO results, due to its ability to bias model weights to zero). While not as severe, the ridge regression-based neural signature contained a lesser degree of sparsity. Realizing this could be a potential signal-to-noise issue, I created two additional versions of the map, smoothed at 2mm and 4mm (fwhm). Subsequent analyses report results using all three of these models. The unsmoothed signature is depicted in Figure 2.4.

[14] It was difficult to conceptualize what a negative association between brain responses and coin flip task values could indicate; in other words, it was difficult to determine whether this kind of brain-behavior relationship was meaningfully bipolar.

Figure 2.4. Custom (sample-specific) neural signature of value.

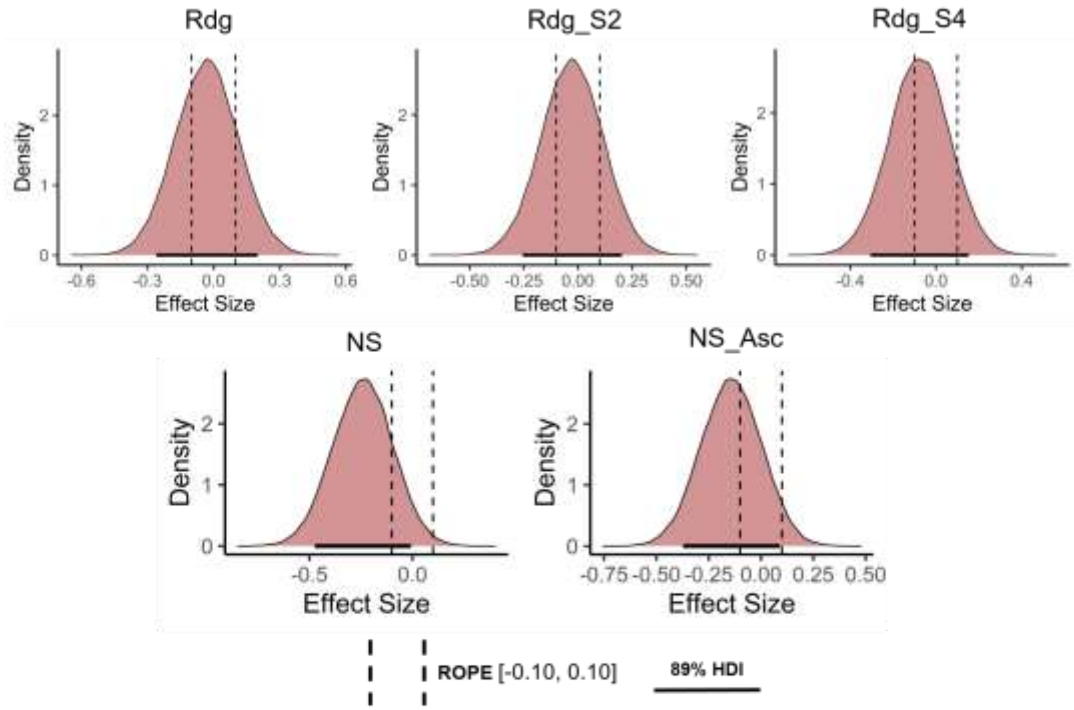


To ensure the custom, sample-specific map was truly indicative of reward, I correlated it with meta-analytic maps of five other constructs: language, pain, working memory, social, and reward (all obtained via Neurosynth). The correlations between the neural signature of value and meta-analytic maps from unrelated constructs (language, pain, working memory, social) were low in magnitude (non-smoothed signature *r*s = -0.068 – 0.001; 2mm-smoothed signature *r*s = -0.063 – 0.018; 4mm-smoothed signature *r*s = -0.052 – 0.019), whereas the correlation between the reward map and the signature was higher (non-smooth signature $r = 0.306$; 2mm-smoothed signature $r = 0.356$; 4mm-smoothed signature $r = 0.505$). This provides discriminant and converging evidence that the signature measures what it is intended to. Further, that the correlation with the meta-analytic map of reward was not very high (e.g., >.7), suggests our signature could be capturing unique or distinct facets of valuation (i.e., it is not redundant with the meta-analytic map). Visual inspection of the signature shows regions canonically associated with reward (e.g., striatum, medial prefrontal cortex) are present in the anticipated direction, further suggesting the signature measures its intended psychological process.

Paired Differences in Value-Based Pattern Expression of Neural Representations. Using the five different neural signatures of value, I observed mixed evidence for the hypothesis that parent and friend neural representations are differentially encoded as a function of value.

Results Using the Sample-Specific Neural Valuation Signature. Results using the sample-specific neural signature of value showed a slight bias towards *friends*, not parents, as indicated by the mean of posterior samples. However, as visualized in Figure 2.4 (top row), roughly equal amounts of the posterior mass lies on either side of ROPE, suggesting the evidence for an effect in either direction is equivocal (*Rdg*: posterior mean, (SD): d = -0.08 (0.14), 89% CI: [-0.31, 0.15]) (*Rdg_S2*: posterior mean, (SD): d = -0.03 (0.14), 89% CI: [-0.25, 0.20]) (*Rdg_S4:* posterior mean, (SD): d = -0.03 (0.14), 89% CI: [-0.27, 0.19]). This outcome was unexpected given the direction of my hypotheses and the results of Study 1. I conducted two *post-hoc*, follow-up analyses to determine whether the results reported here could have been driven by brain regions in the neural signature that were capturing a non-relevant psychological process. While said regions in the signature are, in theory, supposed to be tapping some kind of valuation processes, it is possible that some voxels could be capturing an unrelated or epiphenomenal subprocess, potentially obscuring a signal that would be more consistent with my hypotheses.

Figure 2.5. Posterior distributions of paired differences in value-based pattern expression values.



Note. 'Rdg' and 'NS' refer to the type of signature used (Rdg = sample-specific signature built using ridge regression; NS = Neurosynth signature). 'S' refers to the degree of smoothing in the custom signature (2 = 2mm, 4 = 4mm). 'Asc' refers to the Neurosynth association map (all other maps are uniformity if unmarked). Paired differences are in a standardized metric (*d*). 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals. Difference scores were computed by subtracting friend from parent (parent – friend).

The first *post-hoc* analysis involved excluding primary visual cortex (V1) from the neural signature, under the reasoning that visual processes are unlikely to reflect meaningful information about valuation. While some higher order processes, such as affective salience, have been shown to activate V1, it is not clear that they are psychologically meaningful for the research being conducted here. Re-running the pattern expression analysis with a custom neural signature that excluded V1 voxels did not meaningfully change the results (*Rdg*: posterior mean, (SD): *d* = -0.07 (0.14), 89% CI: [-0.29, 0.17]) (*Rdg_S2*: posterior mean, (SD): *d* = -0.04 (0.14), 89% CI: [-0.27, 0.20]) (*Rdg_S4*: posterior mean, (SD): *d* = -0.04 (0.14), 89% CI: [-0.26, 0.19]).

The second *post-hoc* analysis went a step further and masked voxels in the custom signature that belonged to regions that are thought to share causal relationships with valuation (Dabney et al., 2020; Haber & Knutson, 2009; Kutlu et al., 2021; Lopez-Persem et al., 2020; Vlaev et al., 2011). I again re-ran the pattern expression analysis using only voxels located in the ventral striatum (VS) and medial prefrontal cortex (mPFC). (*Rdg*: posterior mean, (SD): d = 0.07 (0.14), 89% CI: [-0.16, 0.30]) (*Rdg_S2*: posterior mean, (SD): d = 0.02 (0.14), 89% CI: [-0.21, 0.25]) (*Rdg_S4*: posterior mean, (SD): d = 0.02 (0.14), 89% CI: [-0.25, 0.21]). This analysis again yielded equivocal evidence for a value-based signature bias in either direction, even though the sign changed depending on the level of smoothing.

Results Using the Meta-Analytic (Neurosynth) Value Signature. Results using the Neurosynth maps as neural signatures yielded relatively stronger evidence for a value-based bias in friend neural representations (*NS:* posterior mean, (SD): d = -0.24 (0.15) , 89% CI: [-0.48, -0.01]) (*NS_Asc:* posterior mean, (SD): d = -0.14 (0.14) , 89% CI: [-0.37, 0.09]). These findings showed that the majority of the posterior mass either fell within ROPE or in the negatively signed region (which encodes friend > parent). However, once again, these findings suggested that *friend* representations are more strongly encoded as value-based signatures, running contrary what was initially hypothesized (Figure 2.4, bottom row).

Modeling Social Decision Preferences as a Function of Value-Based Pattern Expression. I observed modest evidence to suggest that value-based representations of parents and friend are associated with social decision preferences on the modified CCT. Pattern expression values derived using all three versions of the sample-specific neural value signature predicted social decision preferences (Table 2.1, Figure 2.5).

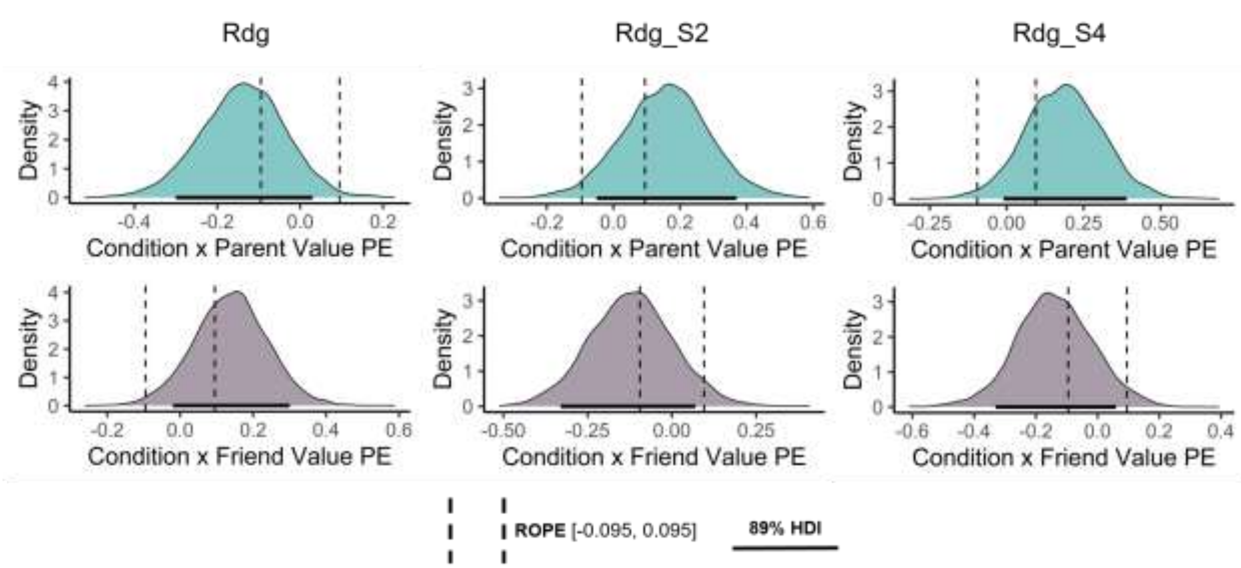
The parameter estimates for interaction terms in the two models using scores from a smoothed neural value signature support the hypothesis that greater value-based pattern representation of a given close other is predictive of favoring said other in the modified CCT. Results obtained from the unsmoothed neural value signature run in the opposite direction (e.g., greater value-based pattern expression of a close other's representation is associated with a *decreased* likelihood of favoring them in the CCT).

Table 2.1. Predicting social decision preferences as a function of value-based representations using a sample-specific neural signature.

Term	Rdg	Rdg_S2	Rdg_S4
Condition	0.30 [0.16, 0.44]	0.30 [0.17, 0.46]	0.30 [0.16, 0.44]
Parent Value PE	-0.02 [-0.23, 0.19]	-0.01 [-0.28, 0.28]	0.01 [-0.24, 0.26]
Friend Value PE	-0.13 [-0.32, 0.08]	-0.01 [-0.26, 0.26]	-0.00 [-0.26, 0.24]
Parent Value PE x Condition	-0.14 [-0.30, 0.03]	0.16 [-0.05, 0.37]	0.19 [-0.01, 0.39]
Friend Value PE x Condition	0.14 [-0.02, 0.30]	-0.12 [-0.33, 0.07]	-0.14 [-0.33, 0.06]

Note. Parameter estimates for the intercept, reward, and risk terms are not reported. 'PE' refers to pattern expression scores, obtained by using each individual subject's parent and friend neural representations and a value-based neural signature. 'Rdg' refers to the type of signature used (Rdg = sample-specific signature built using ridge regression. 'S' refers to the degree of smoothing in the signature (2 = 2mm, 4 = 4mm). Values in brackets represent 89% highest density credible intervals.

Figure 2.6. Posterior distribution plots for model interaction terms capturing the influence of value-based representations on social decision preferences (sample-specific neural signature).



Note. 'Rdg' refers to the type of signature used (Rdg = sample-specific signature built using ridge regression). 'S' refers to the degree of smoothing in the custom signature (2 = 2mm, 4 = 4mm). 'PE' refers to pattern expression score. 'Condition x Parent/Friend' refers to the interaction term entered in the statistical model to assess the association between pattern expression scores and social decision preferences. 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals.

To be consistent with our approach to analyzing paired differences, we conducted two additional *post-hoc* analyses. This involved re-running our hierarchical model (i) with V1 voxels masked out when computing pattern expression scores as well as (ii) computing pattern expression values in reward-related ROIs only. The justification for doing so here is the same as it was for the aforementioned paired differences analysis (e.g., voxels in V1 may not be capturing the targeted psychological process of interest, extending this logic to examine regions that are causally implicated in reward processing). Results for models excluding V1 are listed in Table A2.1; results for models only including reward-related ROIs are listed in Table A2.2.

Overall, the results of these two *post-hoc* analyses are largely consistent with each other, as well as the initial planned analysis: a greater pattern expression score for a given individual was related with a stronger propensity to favor them on the modified CCT. The major difference between them again lies in smoothing of the sample-specific neural signature. The results obtained from using a neural signature that excluded V1 contain the same smoothing-related pattern as the planned analysis—pattern expression scores are directly related to social decision preference when said scores are computed from a smoothed signature (regardless of smoothing magnitude), whereas an inverse relationship is observed when no smoothing is applied to the signature. By contrast, results obtained when using only the reward-related ROIs did not show this pattern—the association between pattern expression scores and social decision preferences are always direct, regardless of smoothing (Visualized in figures A2.1, A2.2).

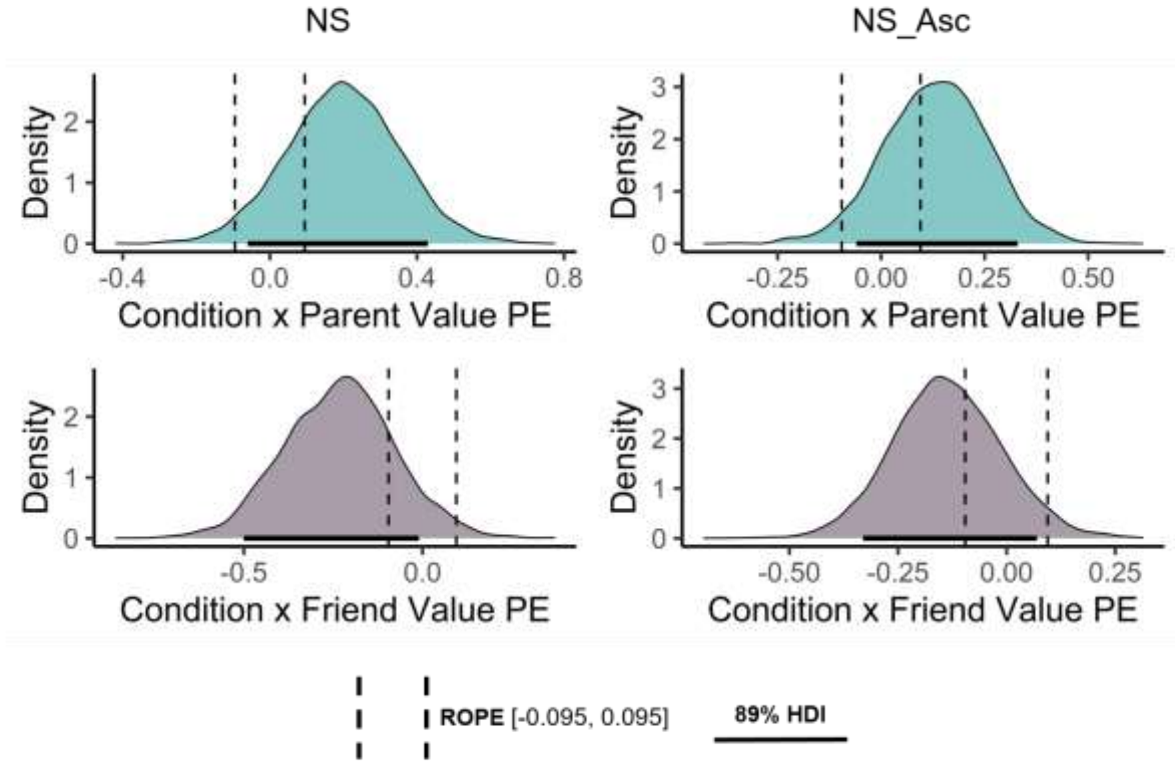
Table 2.2 Predicting social decision preferences as a function of value-based representations using a meta-analytic neural signature.

Term	NS_uni	NS_asc
Condition	0.30 [0.17, 0.44]	0.30 [0.16, 0.45]
Parent Value PE	0.07 [-0.26, 0.40]	-0.14 [-0.38, 0.12]
Friend Value PE	-0.03 [-0.34, 0.30]	0.07 [-0.18, 0.33]
Parent Value PE x Condition	0.20 [-0.06, 0.43]	0.13 [-0.06, 0.33]
Friend Value PE x Condition	-0.23 [-0.50, -0.01]	-0.14 [-0.33, 0.07]

Note. ‘NS’ refers to the type of signature used (NS = Neurosynth signature). ‘PE’ refers to pattern expression score. ‘uni’ refers to uniformity; ‘Asc’ refers to the Neurosynth association. ‘Condition x Parent/Friend’ refers to the interaction term entered in the statistical model to assess the association between pattern expression scores and social decision preferences. Values in brackets represent 89% highest density credible intervals.

Finally, I repeated this analysis using meta-analytically defined neural signatures obtained from Neurosynth (Table 2.2). Conducting the analysis with both the uniformity (forward inference) and association (reverse inference) maps, I once again observed modest evidence to suggest a direct association between pattern expression scores and social decision preference on the CCT (Figure 2.7)

Figure 2.7 Posterior distribution plots for model interaction terms capturing the influence of value-based representations on social decision preferences (meta-analytic neural signature).



Note. 'NS' refers to the type of signature used (NS = Neurosynth signature). 'PE' refers to pattern expression score. 'Asc' refers to the Neurosynth association map (all other maps are uniformity if unmarked). 'PE' refers to pattern expression score. 'Condition x Parent/Friend' refers to the interaction term entered in the statistical model to assess the association between pattern expression scores and social decision preferences. 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals.

Notably, evidence for the effects reported here is modest. Meaning, given the inferential procedure defined in Study 1, I can only rule out one direction for each effect. Using the effect in the top left corner of Figure 2.6 as an example, I can only infer that the parameter estimate is not negative (i.e., the effect is null or positive, meaning that greater parent value-based pattern expression was either related to equivocal preferences or parent preferences, but not friend preferences). While at least a plurality of the posterior mass falls in the direction of the hypothesized effect (parent PE predicting parent preference) for all parameter estimates, a considerable amount also falls within the ROPE and subsequently limits the strength of evidence.

Interim Discussion 2

Study 2 set out to answer one of the biggest questions raised by Study 1: what drives granular social decision preferences among close others? Building upon the prior behavioral approach, this study's use of fMRI and multivariate analysis showed modest evidence that social decision preferences between two close others are driven by the brain expresses these representations in terms of value. These findings carry several implications about the nature of social decision-making.

Excitingly, this study is among the first to examine how neural representations of others influence social decision behavior. Whereas most prior work has examined how individuals process features of each social decision (e.g., the value of each decision alternative, the degree of risk involved, beliefs about a social partner's resources or their attitudes), this study focused on how individuals represent decision partners themselves (specifically, neural representations). This is noteworthy for a few reasons. First, this represents a very direct way to understand the mechanisms that underlie the motivation for one's social decision preferences. Representations of others are, theoretically, all encompassing—they are the lens through which we perceive and contextualize other's behavior (Amodio, 2019; Tamir & Thornton, 2018). Representing one's parent in terms of value would suggest that this close other satisfies or fulfills one's most basic needs and desires (Tottenham, 2020), in turn implying social preferences could emerge from a sense of gratitude or reciprocity to maintain relationship strength. This also signifies that future work should consider investigating whether this is an age-dependent or age-invariant effect. At a *mechanistic* level, these possibilities are more satisfying than correlating relationship quality with decision preference, which arguably provides a description of the phenomenon instead of a mechanism for it. Second, and relatedly, the putative mechanistic influence that representations may have on social decision preference is likely generalizable. In other words, representational influences on decision preferences could be invariant to the decision context because representations of others are theoretically stable and domain-general (though this does not mean that moderating influences may not exist). Third, and by extension, these findings are the first to suggest that a unifying framework for social decision-behavior across contexts could lie with representations of others. Given the putative mechanistic specificity and generalizability of these findings, a focus on representations of others could tie together various forms of social decision-making. Similar endeavors have seen relative success in research involving other constructs (e.g., human perception, memory), underscoring the psychological plausibility of the notion of unifying frameworks (Schurgin et al., 2020; Sims, 2018).

These findings have neuroscientific implications for understanding links between neural value-based circuitry and behavior, as they support a recent trend showing how neural mechanisms thought to process value are also implicated in social cognition and social behavior (Zerubavel et al., 2015). Leveraging the use of representations adds to this line of work by highlighting a potential common thread that links value-based activity across various forms of social behavior. At a broader theoretic level, it should not be surprising to see how the two processes relate to one another: humans are theorized to view others, at least part, in terms of how they satisfy their own individual needs, which in turn motivates social behavior (Amodio, 2019; Tamir & Thornton, 2018). Our results play into this notion, as constructing representations of others using a value-based neural architecture is a seemingly efficient manner of determining the association between other social agents and one’s own goals. Future work may further unpack links between value-based representations and social behavior by integrating information about representations to other value-based frameworks that have recently been used to link to social phenomena (e.g., reinforcement learning and theory of mind).

The lack of consistent evidence for a group-level bias value-based encoding for parent or friend representations in either direction, in conjunction with the presence of modest evidence for individual difference findings, was somewhat surprising. One reason behind this may be due to the fact that parent and friend relationships are highly heterogeneous from person to person, enough so that group-level effects of this sort may be misleading or simply non-informative. The other possibility is that the estimate of the 'true' underlying effect was hampered by virtue of a somewhat small sample size.

Of course, there is a gap between neural representations and the psychology of decision-making. Arguably, a more thorough understanding of the links between representations of others and social decision-making could be achieved by also understanding how *cognitive* representations also contribute to behavior. Examining cognitive representations and links to social decision-behavior has the opportunity to disambiguate the relationship between neural representations and behavior. For instance, even extremely high quality data or methodological sophistication likely never fully guarantee that a researcher has derived one's intended neural signature, or that the components of the neural are psychologically meaningful. Study 3 does just this, using written text data in conjunction with natural language processing to approximate cognitive representations of close others and link them back to decision-behavior.

Study 3: Cognitive Representations and ties to Social Decision-Making

Whereas Study 2 aims to understand the degree to which neural representations of close others affect social decision-making preferences, Study 3 builds upon this work by examining cognitive representations. Cognitive representations are internal models of external stimuli or phenomena that are constructed through semantic labels (e.g., Zemla et al., 2020). Embedded in our brains is a semantic atlas whose labels are manipulated and configured to create working models of the world around us (Huth et al., 2012, 2016; Tversky & Hutchinson, 1986). For every stimulus, it is thought that a subset of semantic labels is cogently bound together to form an emergent network of topics and concepts that constitute a representation. These cognitive representations are deeply important for behavior and the pursuit of social goals (Amodio, 2019; Tamir & Thornton, 2018).

Study 3 used written free-response data in conjunction with natural language processing techniques to determine the degree to which cognitive representations of close others is expressed as a signature of value, conceptually paralleling Study 2. Here I assume that cognitive representations of others can be expressed linguistically and measured through written text (Jackson et al., 2021; Zemla et al., 2020). A corollary to this assumption is that these written terms need to carry meaningful and quantifiable meaning about the structures of these relationships. I specifically used natural language processing (NLP) tools to probe how strongly cognitive representations of participant nominated parents and friends are associated with a topic of value, and relate the strength and direction of such associations to social decision preferences.

A cognitive approach to studying representations is particularly helpful for understanding the *psychology* of social decision-making. An imaging approach, such as the one used in Study 2, is helpful insofar that it can shine light on mental processes that behavioral data alone cannot access. However, solely relying on imaging is inferentially problematic because there is not a clear-cut one-to-one mapping between brain activity and psychological states (Poldrack, 2006). This especially true for an imaging modality such as fMRI, which does not offer the degree of spatial and temporal resolution that would be needed to estimate the neural code underlying particular psychological phenomena of interest. This means that while imaging is successful at capturing a representation, it cannot assess the thematic or psychological content of said representation with complete certainty. This is where a cognitive approach is helpful because it is better suited to measure representational *content*. Pairing an imaging approach with a cognitive approach not only allows for a more comprehensive understanding of how representations guide social decision behavior, but it serves the broader theme of this dissertation which is that in order for social decision-making research to achieve its scientific potential and enhance its external validity, it should strive to be a more personalized science. This involves not only understanding how individuals make decisions involving anonymous others, but how they vary behavior within categories of familiar and close others.

Aside from the relevance of this approach for social decision-making, I argue this study is also beneficial for general research on social representations. While many studies have been devoted to understanding social representations (e.g., Chavez et al., 2017; Guthrie et al., 2021; Wang et al., 2017), very little work actively seeks to tie semantic labels to social representations of specific agents in the way I propose here. This is noteworthy because it promises better mechanistic insight into the nature of associations between representations and related phenomena by virtue of better understanding how specific features of a representation track with the phenomena of interest.

Current Study and Hypotheses. Given previous mean-level preferences towards parents on the social decision-making paradigm, I hypothesize that cognitive representations of parents, relative to friends, will be more strongly linked to a topic of value. Substantively, the logic behind this hypothesis is the same as in Study 2, namely that individuals construct representations of parents, relative to friends, in terms of value because those relationships are better at satisfying internal needs, drives, and desires. However, this study is unique insofar that it assumes value-based features of representations also manifest cognitively as individuals using semantic topics or motifs consistent with value to cognitively construct these representations (Abbott et al., 2012; Jackson et al., 2021; Zemla et al., 2020; Zemla & Austerweil, 2018), and that these are detectable using NLP methods to examine how these labels trickle down into written accounts of relationships with parents and friends. In other words, NLP analyses will help uncover the extent to which cognitive representations of others are expressed in terms of value-based topics or motifs (analogous to the pattern expression analyses in Study 2). Further, I predict that individual differences in these associations will predict social decision preferences such that greater expression of value-based themes in one close other's cognitive representation will be associated with a decision preference for that close other.

Method

Participants. This study aggregated written text data of parent and friend memories that were collected as a salience exercise in Study 1 and other similar studies (Guassi Moreira et al., 2020). Behavioral data from the modified CCT were included in this study to link features of semantic representations to social decision preferences. All data were collected from three independent samples at the University of California, Los Angeles between September 2016 and July 2019. The majority of participants were recruited via the psychology subject pool at UCLA, with a subset (participants administered real rewards from Study 1) being recruited at large from either the Westwood Village area or via instructor announcement at a local community college (Santa Monica Community College). The majority of participants were compensated with course credit whereas those from the subset were paid $20 (USD) for their participation. This resulted in a total sample of 468 young adults. Demographically, 44% of participants self-identified as Asian, 28% identified as White, 3% identified as Black, 2% identified as Native Hawaiian/Pacific Islander, 1% identified as American Indian/Alaska Native, 17% identified as Mixed Race or Other, and 6% declined to self-identify; 21% identified as Hispanic/Latinx; 73% were female. Participants provided written consent in accordance with the policies of the UCLA Institutional Review Board.

Sample Size Considerations. Sample size recommendations for text analysis vary widely between the analysis method (e.g., word2vec embeddings, non-negative matrix factorization, singular value decomposition, etc.) and how it is applied (e.g., Benoit, 2011; Juckett, 2012; Lin & Boutros, 2020). Because this study relied on analysis of existing research records that were initially collected for ancillary purposes, the sample sizes were determined by considering factors unrelated to text analysis. The analysis method for this study, described below, is predicated upon using word embeddings (quantitative representations of words in a latent semantic space) to conduct word embedding association tests. Though the word embeddings came from a pretrained model, thus obviating any sample-size related concerns with training the embedding model, it is unclear precisely how many words are needed to reliably sample participants' representations of their parent and friend and accurately estimate word embedding associations. By aggregating across a sample size of over 450 individuals with at least a few sentences worth of data per individual, I argue that I have approximated meaningful semantic representations of parents and friends for group-level analyses. Additionally, many of the statistical analyses themselves are adequately well powered given they involve relating word embedding associations with trial level decision-making behavior.

Experimental Protocol

Overview. Though the data for this study were taken from several different samples, the protocol was generally consistent across studies: participants nominated a parent and close friend, provided written text data about their relationship with their parent and friend, and then completed a social decision-making task involving hypothetical rewards for each person. These study measures are described in greater detail below.

Text Data Acquisition. Text data were acquired in the same manner as in Study 1. After consenting to participate in the study, participants were told they were going to complete hypothetical decisions on the behalf of a parent and close friend and were prompted to pick one such close other. Participants then filled out a physical document prompting them to recount one memory with each close other (one paragraph suggestions), as well as a handful of words and phrases best describing each other (these data were not used in this study).

Social Decision-Making Paradigm. The modified version of the Columbia Card Task was used to assess social decision preferences. The task design, instructions, and administration were the same as described in prior studies.

Analysis Plan

Overview. The first stage of analyses for Study 3 focused on computing value-based Word Embedding Association Test (WEAT) scores for each participant's written text data. Discussed below, WEAT analyses use word embeddings to compute scores that measure the strength of association between written documents and a given construct of interest. In the context of this study, these scores will serve to quantify how cognitive representations (evaluated via text analysis) of close others are expressed as a function of value. Value-based WEAT scores were subjected to a paired difference analysis—revealing any value-based bias towards parents or friends—and were then related to social decision preferences on the modified CCT. Similar analyses were run on lower level features (length, sentiment) of the written text documents for ‘conceptual parsimony’, testing whether a simpler motivational feature of the data also explain the results.

Calculating Word Embedding Association Test Scores. Word Embedding Association Tests (WEAT) quantify the association between written text documents of interest and a particular topic or construct (Caliskan et al., 2017; Charlesworth et al., 2021; DeFranza et al., 2020; Kurdi et al., 2019). Calculation of WEAT scores rely on word embeddings, vectors of numeric values that quantify the meaning of words. Embeddings are obtained by training an autoencoder-based neural network model to predict the order of words in written text data. While the development of these models was initially spearheaded in industrial settings for text completion on personal devices (e.g., SMS messaging on a smart phone), it was quickly discovered that model parameters carried interesting properties about the semantic meaning of each word (e.g., model parameters for semantically related words are more alike than parameters for semantically unrelated words, such as ‘king’, ‘queen’, and ‘paper’) (Mikolov et al., 2013). Such models involve estimating a weight matrix that effectively maps the position of each word in a corpus to within a latent, *n*-dimensional ‘semantic space’. Each word carries an associated vector that numerically encodes its semantic meaning. Here I used embeddings (vectors) from a popular pre-trained model (the Google News model, *n*-dimensions = 300) that was initially trained and validated on text-scraped Google News data. This implicitly assumes that the semantic representations of words in the current dataset and those from the model are equivalent.

In this study, WEAT scores were calculated at the subject-level by evaluating associations between the terms in parent/friend memories and terms relating to the topic of ‘value’. Consistent with other recent studies (e.g., Kurdi et al., 2019), I began with the term ‘value’ and then searched for closely related synonyms. I only included terms I judged to be most closely related with the meaning of value instead of setting an *a priori* number of words needed for the topic. Items were deemed to be closely related as long as they had some dictionary entry associated with monetary or fiscal value. The final terms are listed in Table 3.1

Table 3.1. Terms related to the topic of value

	Value Terms	
Value	Treasure	Regard
Money	Esteem	Benefit
Prize	Reward	Payment
Appreciate	Valuate	Welfare

Once the terms were identified, value-based WEAT scores were computed for each participant using an iterative procedure. This procedure was performed on one document at a time. For each topic word (*t*), I computed cosine similarity between the embeddings associated the topic word, *t*, and every word, *w*, in the currently targeted document, *d*. Analogous to a correlation, cosine similarity is a widely used way to assess the degree of semantic overlap between two terms (Charlesworth et al., 2021). Cosine similarity scores between *t* and every *w* in *d* were averaged to yield a single value for the *t*. This means that each participant had two sets of topic word scores, one for parent and friend memories, each reflecting the average similarity between each *t* and all words in *d.* In the last step, topic word scores for parent and friend documents were differenced [parent – friend] and the differences were averaged. This yielded a scalar value describing differences in value-based associations between parent and friend memories; a positive score indicates parent memories are more strongly associated with value whereas a negative score indicated friend memories were more strongly associated with value.

WEAT metrics were obtained using the *gensim* package in Python. Stopwords—terms that are ubiquitous in most written documents were removed as a preprocessing step. Words in memories not in the Google News corpus were not included in the analysis.

Extracting Low Level Features of Written Text in Parent and Friend Memories. Two additional features were extracted from the parent and friend memories, length and sentiment, with the goal of gauging whether basic features of memories carry any motivational significance. Both metrics were obtained using the *TextBlob* Python package. Memory length was defined as the number of words comprising a memory. Memory was extracted for each memory by using the package's sentiment function. The function assesses the sentiment—termed polarity, in the package's parlance—of a given piece of text by using model weights from a Naïve Bayes classifier that was pre-trained on a corpus of positive and negative movie reviews. Sentiment scores for each memory were bound between -1 (negative) and 1 (positive). Positive and negative in this sense are defined by human affective valence (not necessarily in terms of value). Put plainly, polarity is how emotionally 'good' or 'bad' a piece of text is. Scores were created for both of these metrics in the same manner as WEAT scores (parent – friend)

Statistical Analysis. I conducted two sets of analyses on the written text data, mirroring the analytic structure in Study 2: examining paired differences on written text metrics and relating said text metrics to social decision preferences.

I once again analyzed paired differences in metrics from the written text data in the same Bayesian framework as Study 2. This analysis was repeated three times: once for differences in parent and friend memory length, another for differences in parent and friend memory polarity, and a third time for differences in value WEAT scores[15].

The second step tested whether individual differences in written text metrics predicted social decision preferences using the same hierarchical Bayesian logistic regression framework described in Studies 1 and 2. The form of these models in this study is identical to the prior two studies. Two models were run, one model with lengths and polarity values for parent and friend memories predicting social decision preferences (4 between-person predictors), and a second model where value-based WEAT scores predicted such preferences (1 between person predictors).

For all analyses here, the inferential criteria were the same as in Study 1. For paired analyses, the ROPE was expanded slightly to [-.1 to .1], as effects less than 1/10th of one standard deviation were of no interest.

Results

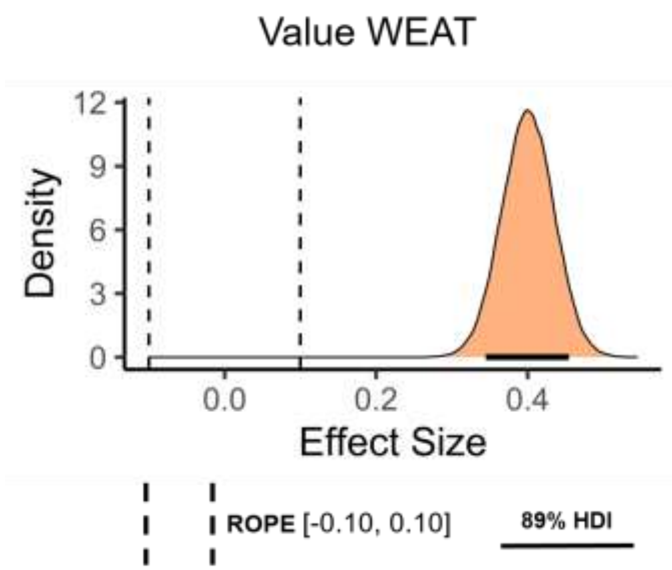
Descriptive Statistics. Descriptive statistics were computed for metrics of interest. The mean length for parent memories was 20.59 words (9.30 SD); the mean memory length for friend memories was 19.15 words (8.74 SD). Polarity (sentiment) for both parent and friend memories were slightly positive (mean parent memory polarity: 0.15 (0.26 SD); Mean friend memory polarity: 0.13 (0.26 SD)). Finally, the mean of the value-based WEAT scores was 0.006 (0.015 SD). Recall that an average greater than zero would indicate parent memories are more

[15] While the WEAT analysis may seem more analogous to a one sample t-test in a Bayesian framework since only a single vector of word embedding association scores are used, we remind the reader that these values are already difference scores between parent and friend word embedding associations.

strongly associated with the value-based topic whereas an average less than zero would indicate friend memories are more strongly associated with the topic.

Paired Differences in Written Text Metrics. Bayesian paired difference analyses suggested meaningful differences in the structure and content of the parent and friend memories. Foremost, results using the value-based WEAT scores show that parent memories are more strongly associated with the topic of value than friend memories (posterior mean, (SD): d = 0.40 (0.03), 89% CI: [0.34, 0.45]; Figure 3.1).

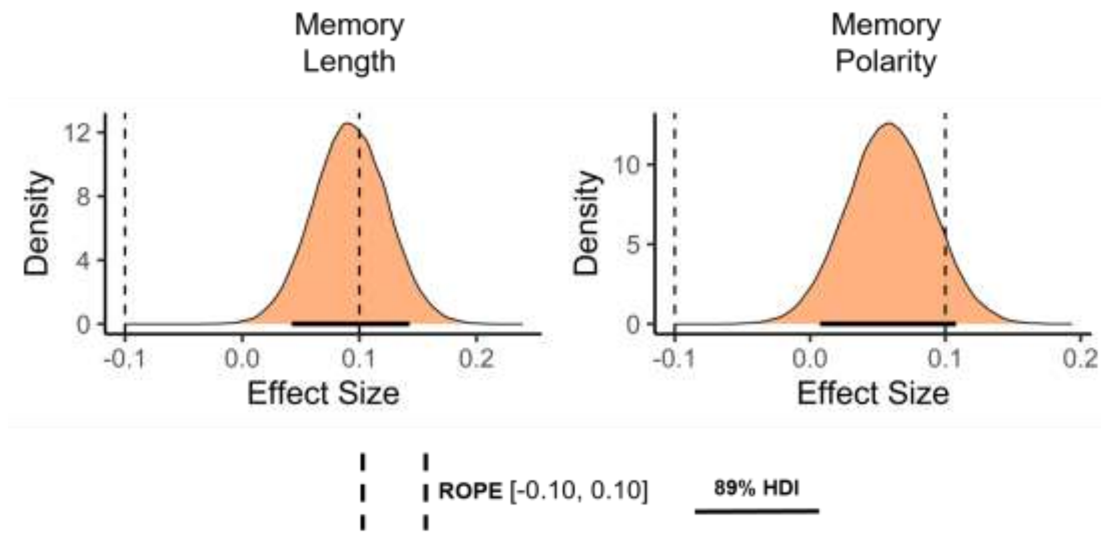
Figure 3.1 Posterior distributions of paired differences in value-based WEAT scores.



Note. Paired differences are in a standardized metric (d). 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals. 'WEAT' refers to Word Embedding Association Test.

The results for memory length and polarity were not as robust. The average effect sizes suggest parent, relative to friend, memories are slightly longer (in terms of words) and slightly more positive in their sentiment. However, because part of the posterior mass for these effects fall within ROPE, I can only conclude that memories are not longer or more positive for friends (i.e., I can only rule out the sign is not negative, but the true effect could be zero or positive) (memory length differences: posterior mean, (SD): d = 0.09 (0.03), 89% CI: [0.04, 0.14]; polarity differences: posterior mean, (SD): d = 0.06 (0.03), 89% CI: [0.01, 0.11]; Figure 3.2).

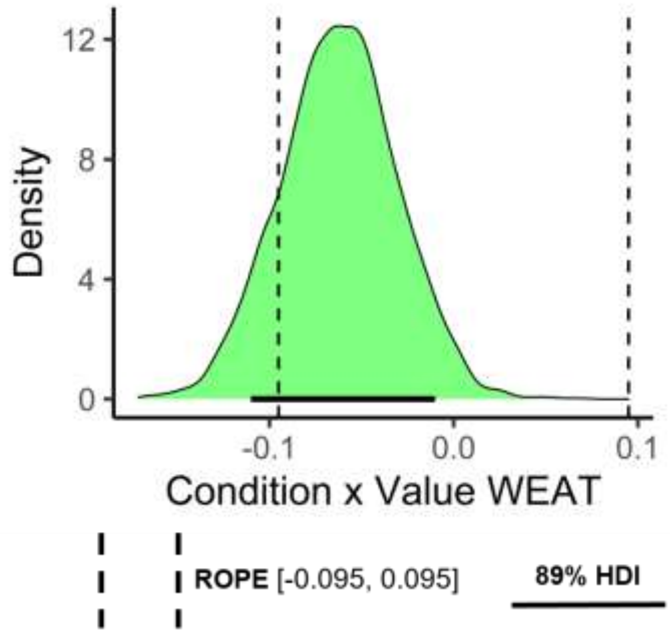
Figure 3.2 Posterior distributions of paired differences in memory lengths and sentiment.



Note. Paired differences are in a standardized metric (d). 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals. Memory length was measured as the number of words; 'polarity' is a measure of sentiment.

Modeling Social Decision Preferences as a Function of Written Text Metrics. Results relating value-based WEAT scores to social decision preferences ran contrary to predictions, with the sign of the relevant parameter estimate being *negative* and indicating that stronger expression of value-based terms in parent memories was associated with a greater propensity to favor one's friend (and similarly, stronger expression of such terms in friend memories was associated with a greater propensity to favor one's friend (Figure 3.3; Table 3.2).

Figure 3.3. Posterior distribution model interaction term capturing influence of value-based WEAT scores on social decision preferences.



Note. 'Condition x Value WEAT' refers to the interaction term entered in the statistical model to assess the association between value-based WEAT scores and social decision preferences. 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals. 'WEAT' refers to Word Embedding Association Test.

Table 3.2. Predicting social decision preferences as a function of value-based WEAT scores.

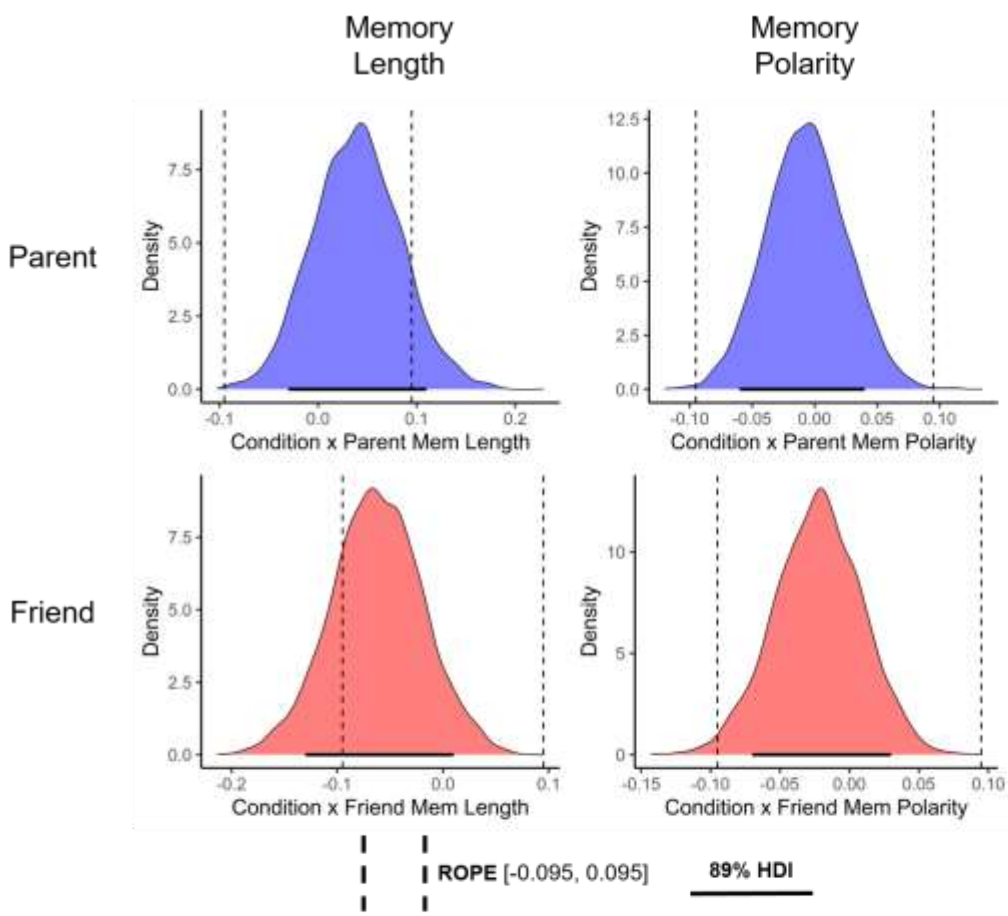
Term	
Condition	0.26 [0.21, 0.31]
Value WEAT	0.06 [0.00, 0.12]
Value WEAT x Condition	-0.06 [-0.11, -0.01]

Note. 'WEAT' refers to Word Embedding Association Test. 'Condition x Value WEAT' refers to the interaction term entered in the statistical model to assess the association between value-based WEAT scores and social decision preferences. Values in brackets represent 89% highest density credible intervals.

Most of the posterior mass of this effect fell on the negative side of the parameter space and ROPE, meaning that we can only have evidence to strictly conclude the effect is not positive.

Parallel findings with the memory length and sentiment metrics are shown in Table 3.3 and Figure 3.4. These effects show modest evidence that memory length serves as a motivational index of social decision behaviors, as having longer memories for a given close other was associated with a slightly stronger preference to favor them during the modified CCT. By contrast, there was a null effect with sentiment, suggesting the effect is practically equivalent to zero.

Figure 3.4 Posterior distribution model interaction term capturing influence of memory length and sentiment on social decision preferences.



Note. 'Condition x [Text Feature]' refers to the interaction term entered in the statistical model to assess the association between a given feature of written text and social decision preferences. 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals. 'Mem Polarity' refers to memory sentiment; 'Mem Length' refers to memory length (number of words).

Table 3.3 Predicting social decision preferences as a function of memory length and sentiment.

Term	
Condition	0.26 [0.21, 0.31]
Parent Memory Length	0.05 [-0.04, 0.13]
Friend Memory Length	-0.04 [-0.13, 0.03]
Parent Memory Polarity	0.00 [-0.06, 0.07]
Friend Memory Polarity	-0.01 [-0.07, 0.05]
Parent Memory Length x Condition	0.04 [-0.03, 0.11]
Friend Memory Length x Condition	-0.06 [-0.13, 0.01]
Parent Memory Polarity x Condition	-0.01 [-0.06, 0.04]
Friend Memory Polarity x Condition	-0.02 [-0.07, 0.03]

Note. 'Condition x [Text Feature]' refers to the interaction term entered in the statistical model to assess the association between a given feature of written text and social decision preferences Values in brackets represent 89% highest density credible intervals.

Interim Discussion 3

Study 3 aimed to examine representations of parents and friends at the cognitive level, using topic-focused NLP analyses to supplement and disambiguate the relationship between value-based processes, representations, and social decision preferences. Using written memories of parents and friends, I observed evidence to suggest that written accounts of close others contain motivational information that predict social decision preference, seemingly in a value-based fashion. However, the direction of these effects were mixed, requiring additional unpacking by future studies. These findings, and their implication, are discussed in greater detail below.

At the most basic level, results from this study suggest that written accounts of close other contain information about (i) representations of close others and (ii) relatedly, motivational processes that influence social decision preferences. Two pieces of evidence support these claims. First, there was a robust group level-effect for differences in parent and friend memories in value-based WEAT scores, suggesting that parent memories are more strongly encoded in terms of value. Consistent with my hypothesis, this finding mirrors group-level effects in behavior that show individuals consistently favor a parent over a friend during social decision-making. A conceptually similar, albeit less robust, finding emerged with sentiment (mean level parent preference). However, the fact that results with sentiment were less robust underscores the specificity of the WEAT analysis, the latter is not synonymous with the former. While this effect was inconsistent with the analogous analysis involving neural representations, the NLP results broadly suggest that written text contains meaningful information about representations of close relationships, partly validating the approach used here and potentially aiding future similar research. Second, written text metrics were predictive of social decision behavior indicating that written documents carry motivational information. Longer memory lengths for a particular close other were associated with a greater tendency to favor said close other, whereas value-based WEAT scores were inversely associated with social decision preference. The latter effect is somewhat surprising, as it runs contrary to what was hypothesized, as well as what was observed with the analogous analysis in Study 2. One explanation for this could be that the two methods of assessing value (pattern expression, WEAT) are tapping different aspects of the construct. Indeed, many psychological constructs are multifaceted and high dimensional (Caspi et al., 2005; Robins et al., 2001), so perhaps it is unsurprising that value is a multifaceted construct that motivates behavior in inconsistent ways. Following this logic, it is apparent that psychological science may have more work to do in terms of accurately defining value from a psychological perspective using empirical methods. Although scholars have written at length about what value means from a psychological perspective, experimental work seems to favor using monetary value because it is easy to operationalize and often represents a known input for modeling purposes. While experimentally convenient, it sidesteps the issue of comprehensively defining the psychology of ‘general’ value from an empirical perspective.

An alternative possibility could be due to the spontaneous versus effortful nature of the two assessments. Spontaneous representations of parents and friends were measured in Study 2, theoretically capturing implicit aspects of the representation. By contrast, deliberate and effortful responses were required to measure the representations here in Study 3. These differences could mean that different facets of parent and friend representations were measured, subsequently having different consequences for behavior. This could have consequences for motivations in social decision-making—the spontaneous facet of representation may capture a baseline, or fundamental preference whereas a more deliberate or effortful facet may open the door for the influence of other top-down goals (such as motivation to repair a relationship or devote more time to a stable one). Yet another explanation could be due to the materials used to derive neural or semantic signatures of value. Monetary stimuli may tap a more ‘global’ value signature in the brain—evidenced by the fact that neural signals of social and monetary value coarsely share similar patterns of brain activity (Wake & Izuma, 2017)—whereas they may be more narrowly tapping strictly financial value in the linguistic domain. Greater value-based cognitive representations in Study 3 could perhaps reflect financial security in a relationship, and participants may have been thus more motivated to acquire resources for friends who have less of these resources compared to parents. One future direction for this type of work would involve a more comprehensive linguistic assessment of value. However, defining and assessing ‘social value’ figures to be difficult because non-monetary terms associated with social value are likely also strongly related to other psychological constructs (Seaman et al., 2016). Future work that attempts to understand representational structure in terms of social value will therefore likely need additional studies to lay the groundwork to precisely define unique linguistic markers of social value (e.g., Rhoads et al., 2021).

A complicating factor for this study may be in the length of the parent and friend memories. Based on available research, it is unclear precisely how many words per document are needed to extract meaningful estimates of one's desired metrics. Although the number of participants was relatively large in this study, and we had additional power for individual difference analyses given the high volume of trial-level decision-making data, the stability of polarity and value-based WEAT scores is still up for debate. Perhaps it is possible that magnitude, precision (posterior width), or sign of the observed effects would change with additional data from the written memories. Moreover, it is also unclear how well the written memories adequately sampled one's cognitive representation. These memories may have captured too much or noise, or state-induced variability to be called a truly 'global' estimate of the representations of interest. Perhaps this is an area where future studies leveraging ecological momentary assessments or high density sampling can be of use (Poldrack et al., 2015; Salehi et al., 2020), if not to at least further gauge the feasibility of the approach described here for future research.

These results also contribute to other broader literatures. First, they speak to the importance of using linguistic data as a window to studying psychological processes (Jackson et al., 2021; Kurdi et al., 2019). These results further affirm that meaningful information about psychological processes (e.g., representational content, implicit biases) are embedded in linguistic data, and studying linguistic data is one way of extracting insights about said processes. Second, these results also support recent literature showing that cultural information and norms are reflected in linguistic output (Arseniev-Koehler et al., 2021; Arseniev-Koehler &

Foster, 2020; Charlesworth et al., 2021), to the extent that social decision-preferences reflect cultural information. Both of these links to broader research support the argument that increasing the external validity of social decision-making research will likely help this research serve and inform other arcs of research in psychology and related sciences.

General Discussion

Overview.

The purpose of this dissertation was to address the issue of external validity in social decision-making by pushing this research towards a personalized science. Because such research almost exclusively focused on unfamiliar others, I argued that many theoretical motifs derived from prior social decision-making research are not actually meaningful at describing most real-world decisions, since such decisions involve close others as well as the need to grapple with conflicting outcomes. Using late adolescents as a model population, this dissertation demonstrated that individuals evinced robust and consistent parent-over-friend preferences, and that these preferences may be driven by value-based expression of close other representations (inferred from individual differences in value-based representations and social decision behavior). I discuss the implications of these findings for developmental science, social neuroscience, social decision-making, value-based processes and social interactions, and the study of social behaviors, more generally. I conclude by enumerating limitations and considering future directions.

Social Decision-Making and Developmental Science.

In the introduction I noted that developmental science stood to gain by studying social decision preferences in late adolescence. I argue that the results here lend support to this notion because the information that late adolescents tend to favor their parents over friends informs the field by defying traditional notions of adolescence in two ways (Steinberg & Morris, 2001). First, it shows that parents are still central figures in the lives of their adolescent offspring (Crone & Fuligni, 2020). While individuals become increasingly oriented towards peers in adolescence (Blakemore & Mills, 2014), this dissertation complements recent work highlighting the importance of parents to adolescents (Telzer et al., 2015) by showing that teens prefer their parents *over* their friends in some instances. This is particularly important and novel because the field has never pitted adolescent and parent outcomes in this way, thus revealing the relative importance of parents and friends in one particular context. Second, this work adds to the growing body of literature that adolescents are not indiscriminately prone to taking risks regardless of context, instead lending support to the broad notion that adolescents titrate risky behavior in accordance with socioemotional goals (Pfeifer & Berkman, 2018).

Because human beings develop in dynamic ways, it will be interesting to repeat this work in other age groups and with different types of social agents or close others. The behavioral and analytic paradigms introduced here carry potential value for systematically mapping social decision preferences among an individual's set of close others across development. In thinking about what future studies may reveal, it is an open question whether the same pattern of findings would be observed in other age groups, such as early adolescence, for instance, when individuals are first beginning to show heighted peer orientation (Ahmed et al., 2020; Foulkes et al., 2018). Relatedly, while it is likely that preferences shift across development, it is unclear how they do so (e.g., do individuals prefer parents as children, friends as early adolescents, parents as late adolescents and young adults, romantic partners or their own children in middle adulthood?). The current set of results will likely inform work pursuing these aforementioned novel avenues.

Implications for Social Neuroscience.

The neuroscience findings here are interesting because they suggest the brain isn't simply relying on two or three node circuits to perform low dimensional computations over decision-level inputs during social decision-making (e.g., computing subject value of a safe or risky option based on the degree of reward, uncertainty, etc.) (Gangopadhyay et al., 2021; Rilling & Sanfey, 2011). Instead, the results of this dissertation indicate that the brain is likely performing a multitude of high dimensional calculations using finely encoded representational information about others to guide decision behaviors. This is consistent with the notion that representations themselves are intrinsically high dimensional, given that they require storing and integrating a wealth of information in order to make real time predictions (Kriegeskorte & Douglas, 2018). However, because I only measured one specific feature of these representations (value expression), there remains more to be known about the mechanistic details of these high dimensional computations. A more integrative social neuroscience may help address this by leveraging animal models of social decision-making (Ben-Ami Bartal et al., 2011; Dal Monte et al., 2020) to better decode the specific computations by which representations guide behavior.

This work also contributes to the growing literature that social processes, generally speaking, are value based (Hackel & Zaki, 2018; Zerubavel et al., 2015). By this I mean that value is used as a heuristic for tracking, integrating, and acting upon social information. In the case of this dissertation, value-based encoding of social agents was related to actionable decision preferences for said individuals, encouraging social and affective neuroscience to study personalized social behavior in the context of value-based behavior. Tools in fields such as computational science could be leveraged to understand how the brain implements basic social cognitive skills in a value-base manner, as well as how social contexts and top-down preferences modulate these processes.

What Do These Results Tell Us About Social Decision-Making?

Psychological science has cared about social decision-making for all of this century and most of the last, even if the term 'social decision-making' was less ubiquitous then as it is now. Throughout this time, social decision-making research has been subject to many of the same goals that drive research elsewhere in psychological science: to form comprehensive theories of behavior as it occurs in the real world. While the extant literature on social decision-making is detailed enough to allow for synthesis of some broad and general motifs that color social decision behavior (Feldmanhall & Chang, 2018; Sokol-Hessner & Rutledge, 2019), they were subject to a critical caveat: it was unknown how social decisions varied as function of social target. The findings in this dissertation suggest that unique agents within a social category (e.g., close others) can be the subject of specific and consistent social decision preferences and thus that several motifs thought to generalize across social decision-making research, in fact do not (Carlson et al., 2020; Crockett et al., 2017; Feldmanhall & Chang, 2018). For instance, harm prevention—or aversion to actions that would inflict negative outcomes on others—is often assumed to be a motivational staple to social decision making. Yet, if harm prevention were truly a ubiquitous motivation in social decision-making research, individuals in the present studies would have chosen equally or randomly between parents and friends, minimizing the collective 'harm' for each close other. Instead, individuals clearly prioritized one individual over another, showing that the notion of harm prevention in social decision-making is modulated by other motivations. Similarly, individuals are thought to minimize uncertainty in their environments (FeldmanHall & Shenhav, 2018), either by avoiding situations containing risk. Again, the results in the current dissertation dispute this theme, as individuals explicitly sought out riskier choice scenarios if meant benefitting a preferred individual.

These results underscore that social decision-making is exquisitely dependent on the target of said decision-making and that conclusions drawn exclusively from decisions about distant others are unlikely to generalize to decisions about close others. It is clear that there are other cognitive and affective mechanisms that modulate social decision processes involving computation of harm, uncertainty, etc. More information is needed to systematically map the factors that cause individuals to exhibit varied prioritizations of different individuals (e.g., decomposing associations between individual differences in relationship quality and social decision preferences). Future work should consider adopting more formalized models of constructs like relationship quality, and relate individual-level model parameters to social decision parameters as a way to understand how specific computations become co-opted to fulfill motivational goals associated with a particular relationship. Identification of these factors should then hypothetically allow for a more appropriate integration of other computations (such as harm prevention, uncertainty minimization, etc.), as these latter computations are likely extensions or implementations of more sweeping motivational forces.

What Do These Results Say About Value and Social Interactions?

This dissertation found repeated, if modest, evidence that value-based information underlies social representations and guides social decisions. These findings are consistent with the broader literature insofar that they also find value-based processes are applicable in social contexts (Hackel et al., 2017; Pärnamets et al., 2019; Zerubavel et al., 2015). While prior work indeed alluded to the possibility of value-based representations (e.g., Hackel et al., 2017), we did not know whether or how this manifested at an individual level. The findings from these studies are thus novel because they can directly speak to how representations of *specific* individuals are expressed in terms of value, unlike prior studies that simply showed value-based processes or value-based neural circuitry tracked information about social structures, or generic groups of individuals.

These findings highlight the importance of value-based processes in social behavior, and nudge social scientists to abandon traditional conceptualizations of 'value' being solely rooted in a purely monetary construct (Wake & Izuma, 2017). Put differently, they highlight that value-processes may be so important for guiding behavior, that we build representations of others based on how much these others support us. Instead, we must embrace value-based processes as drives that optimize the acquisition of resources that fulfill our needs and desires. While monetary operationalizations of value are convenient, they may not fully capture the entire psychological substance of value. In fact, I would go so far to argue that value-based theories are simply formalized re-conceptualizations of several founding tenets of psychological science (e.g., law of effect), meaning that there is already a framework in place by which to update and revise psychological conceptions of value Concretely, I believe this endeavor may be carried out the way I described in Interim Discussion 3 (empirically identifying different forms of value that are unique and distinct from other constructs).

Implications for the Study of Social Behavior and the Future of Psychological Science.

The findings described in this dissertation have relevance for other arenas within psychology that concern themselves with social behavior (e.g., social psychology, health psychology, clinical psychology). In particular, my findings regarding close others (in contrast to prior research on unfamiliar others) underscore recent recommendations to increase the adoption of naturalistic study designs (DuPre et al., 2020; Finn et al., 2020; Grall & Finn, 2021; Jolly &

Chang, 2019). This is because the central theme of this dissertation's results (*within*-category granularity of social behavior involving close others) could conceivably apply to any kind of psychological process that involve unfamiliar, familiar, and close others. That I observed granularity within the category of close others merits a consideration of how this could apply to other research topics. Take for instance studies of intergroup dynamics, which routinely find that social perception of individuals varies by group membership (ingroup vs outgroup) (Hackel et al., 2017; Van Bavel et al., 2008). Because social categories are overlapping and nested, it is plausible differences could emerge as one superimposes additional categories within a broad ingroup (familiar others could be broken down into groups based on friends, family, close relationships, distant relationships, etc.) (Lockwood et al., 2021). Research topics such as empathy and theory of mind could also be impacted, as the results presented herein are seemingly a function of motivational factors, signifying there is no compelling reason to prematurely conclude such motivational factors will not influence social phenomena investigated by various other research topics. Continuing to use laboratory or observational paradigms that solely rely on unfamiliar others could be problematic insofar that such approaches result in low-dimensional distillations that inadequately approximate their intended targets in the real-world and therefore sacrifice generalizability and external validity of subsequent results.

Extrapolating this line of reasoning to its logical conclusion, one could argue that the future of psychological science—at least as it relates to understanding social behavior—lies in personalized science. Human processes are high dimensional and emergent, and thus theories must incorporate this dimensionality in order to be useful at making scalable, generalizable, and unintuitive predictions in the real world (Jolly & Chang, 2019). Personalization accommodates high dimensionality and could open new doors for psychological scientists of the future, in both basic and applied settings. Pursuing studies of this ilk will be challenging, as it will likely necessitate greater reliance on team science (e.g., Bouwmeester et al., 2017; Hagger et al., 2016; McCarthy et al., 2018; Verschuere et al., 2018) as well as collaborations with private businesses, both of which figure to have the infrastructure to collect high volume data needed to reliably estimate the highly parameterized models that are required in large scale personalized science. It will also require novel methodologies, as current social decision-making research disproportionately uses laboratory based tasks to examine social decision tendencies. Such development of novel tools and measurement approaches to study social decision behavior as it occurs outside the laboratory will help enhance the precision and generalizability of social decision research. Such improvements could involve the incorporation of existing methods, such as scraping social media for naturalistic information (e.g., Lindström et al., 2021) and leveraging the use of ecological momentary self-reports, or developing novel ways to passively measure information about social decisions (e.g., an app that automatically detects self-disclosed social decisions based on text message communications). While there is much work to be done, I believe these humble findings join the chorus of constructive critiques aimed at improving the rigor of psychological science in the domain of social decision-making.

Limitations and Future Directions.

There are several limitations to the current dissertation, each falling into one of three broad following categories: sampling and assessment of social decision preferences, statistical modeling, and measurement of representations.

The most pressing issue in the first category (sampling and assessment of social decision preferences) concerns generalizability (Simons et al., 2017). Because participants were exclusively recruited and tested in the western United States, it is possible that a different pattern of results would be observed elsewhere because of cultural differences surrounding parent–child relationships. Further, our procedure involved each participant nominating just one parent and one friend, leaving open the possibility that results could differ with different close others owing to qualitative differences between different relationship partners within a particular category. Another issue to consider is the use of monetary rewards, which may be marked by cohort effects due to the current economic climate. Related to assessment of social decision tendencies, it is possible that the parent-over-friend preferences observed in the modified CCT across all three studies were, at least partially, a product of social desirability (Furnham, 1986; Phillips & Clancy, 1972). The most direct way to address this issue is to consider replicating Study 1 while statistically adjusting for social desirability. With that said, I argue the likelihood that social desirability confounded the results is minimal on several grounds. First, social desirability is known to be trait-like, varying widely between individuals (Phillips & Clancy, 1972). I have no reason to suspect I inadvertently oversampled individuals on the basis of social desirability (and at that point, it is possible any potential such confound could actually be caused by other correlated traits). Second, my experiment contained features that have long been known to reduce susceptibility to social desirability confounds, such as the lack of a clear, socially desirable outcome[16] and self-administration (the experimenter unobtrusively monitored participants and no other individuals were present) (Furnham, 1986). Moreover, I *did* observe that social decision preferences were explained a function of several other between-person variables. If social desirability were *causing* the results observed here, it would imply (though

[16] Would a majority of participants unequivocally think it is more socially desirable to favor a parent? If so, wouldn't this also be meaningful in its own right? Regardless, if the desirability of both options (favoring parent versus favoring friend) is roughly equivocal, then individual differences in social desirability is likely not a confound.

not guarantee) that the other predictive features of social decision preferences (neural and cognitive representations, relationship quality) would be correlated with social desirability in order evince a statistical relationship with decision preferences. I find that to be unlikely. Nevertheless, this consideration underscores the importance of identifying dispositional influences on targeted social decision behavior.

Related to the second category of limitations (statistical modeling), it is clear that future work must focus on understanding how flexibility in the statistical quantification of social decision preferences affects results, particularly when using highly flexible Bayesian analytic techniques. Recent studies have emphasized how analytic flexibility can drastically change the results and conclusions of a study (Botvinik-Nezer et al., 2020; Silberzahn et al., 2018). As a result of this, it is now recommended that individuals run their analyses with many, if not all, conceivably defensible pipelines and track the robustness of ensuing results. While this approach is outside the scope of the current dissertation, it is nevertheless a worthwhile endeavor for future work. Relatedly, it remains to be seen how sensitive social decision studies like the one presented here are to differences prior specifications for Bayesian analyses. Future studies could likely gauge this via extensive, albeit computationally costly, simulation studies.

The third category of limitations concerns measure of neural and cognitive representations. The most notable issue in this category is one of sample size. Namely, did I collect enough data, both within- and between-subjects, to properly estimate neural and cognitive representations of close others? While I did the best within my means to address potential sample size issues, it is now known that sometimes tremendous amounts of data are required to estimate personalized data, especially in imaging modalities (Helmer et al., 2020; Marek et al., 2020; Marek & Dosenbach, 2018). It remains to be seen what effect this had on the current results, as it is unclear whether potentially limited sample sizes would affect the width of the posterior distribution, the magnitude of the posterior mean, the sign of the posterior mean, or all three. This issue may be the root cause of the modest evidence observed in Studies 2 and 3, which was not robust enough to definitely conclude the effect of interest lied in one particular direction (instead, it was only strong enough to rule out the direction opposite the sign, leaving me unable to rule out a null finding or an effect in the direction of the sign). To avoid these pitfalls, future studies may consider ‘high density sampling’ techniques that acquire many data from a relatively limited subset of participants. In this context, this could entail recruiting subjects for several months (e.g., 3-5) of ecological momentary assessments that use written text to record the nature of daily interactions with parents and friends and using the ensuing rich dataset to derive more accurate cognitive representations of parents and friends. In an imaging context, such an approach could involve preselecting subsets of individuals (e.g., 5 participants/subset) with strong parent and friend preferences and administering several hours worth of personalized parent and friend stimuli (perhaps in a manner more engaging than the one used here in Study 2) for hours (e.g., 7-10 hours) to fully tap neural representations of parents and friends (Huth et al., 2016) and then compare differences in said representations.

Conclusions.

In summary, social-decision making preferences systematically vary as a function of whom is affect, and late adolescents transitioning to young adulthood tend to favor their parents over friends overall (Study 1). Considerable inter-individual heterogeneity in these preferences were explained by the degree to which representations of close others were encoded as value-based computations, both at the neural (Study 2) and cognitive levels (Study 3). It is my hope that these results help push social decision-making research towards greater external validity.

Appendix A

Supplemental Results

Table A2.1. Predicting social decision preferences as a function of value-based representations using a sample-specific neural signature, excluding primary visual cortex (V1).

Term	Rdg_Sig	Rdg_S2_Sig	Rdg_S4_Sig
Condition	0.29 [0.16, 0.43]	0.30 [0.15, 0.44]	0.30 [0.15, 0.44]
Parent Value PE	-0.03 [-0.22, 0.17]	-0.11 [-0.33, 0.12]	-0.11 [-0.33, 0.12]
Friend Value PE	-0.14 [-0.35, 0.05]	-0.02 [-0.18, 0.25]	0.02 [-0.21, 0.24]
Parent Value PE x Condition	-0.14 [-0.28, 0.01]	0.05 [-0.13, 0.22]	0.11 [-0.05, 0.30]
Friend Value PE x Condition	0.14 [-0.02, 0.29]	-0.01 [-0.17, 0.17]	-0.04 [-0.22, 0.13]

Note. Parameter estimates for the intercept, reward, and risk terms are not reported. 'PE' refers to pattern expression scores, obtained by using each individual subject's parent and friend neural representations and a value-based neural signature. 'Rdg' refers to the type of signature used (Rdg = sample-specific signature built using ridge regression. 'S' refers to the degree of smoothing in the signature (2 = 2mm, 4 = 4mm). Values in brackets represent 89% highest density credible intervals.

Table A2.2. Predicting social decision preferences as a function of value-based representations using a sample-specific neural signature, including only reward regions (VS, mPFC).

Term	Rdg_Sig	Rdg_S2_Sig	Rdg_S4_Sig
Condition	0.30 [0.16, 0.44]	0.30 [0.16, 0.44]	0.30 [0.16, 0.44]
Parent Value PE	-0.03 [-0.27, 0.21]	-0.00 [-0.23, 0.24]	-0.03 [-0.28, 0.20]
Friend Value PE	-0.04 [-0.27, 0.19]	-0.05 [-0.28, 0.18]	-0.02 [-0.25, 0.23]
Parent Value PE x Condition	0.18 [0.00, 0.36]	0.14 [-0.05, 0.32]	0.14 [-0.05, 0.33]
Friend Value PE x Condition	-0.11 [-0.28, 0.08]	-0.10 [-0.28, 0.08]	-0.09 [-0.28, 0.08]

Note. Parameter estimates for the intercept, reward, and risk terms are not reported. 'PE' refers to pattern expression scores, obtained by using each individual subject's parent and friend neural representations and a value-based neural signature. 'Rdg' refers to the type of signature used (Rdg = sample-specific signature built using ridge regression. 'S' refers to the degree of smoothing in the signature (2 = 2mm, 4 = 4mm). Values in brackets represent 89% highest density credible intervals. 'VS' refers to ventral striatum, mPFC refers to medial prefrontal cortex.

Figure A2.1. Posterior distribution plots for model interaction terms capturing the influence of value-based representations on social decision preferences (sample-specific neural signature, excluding V1).

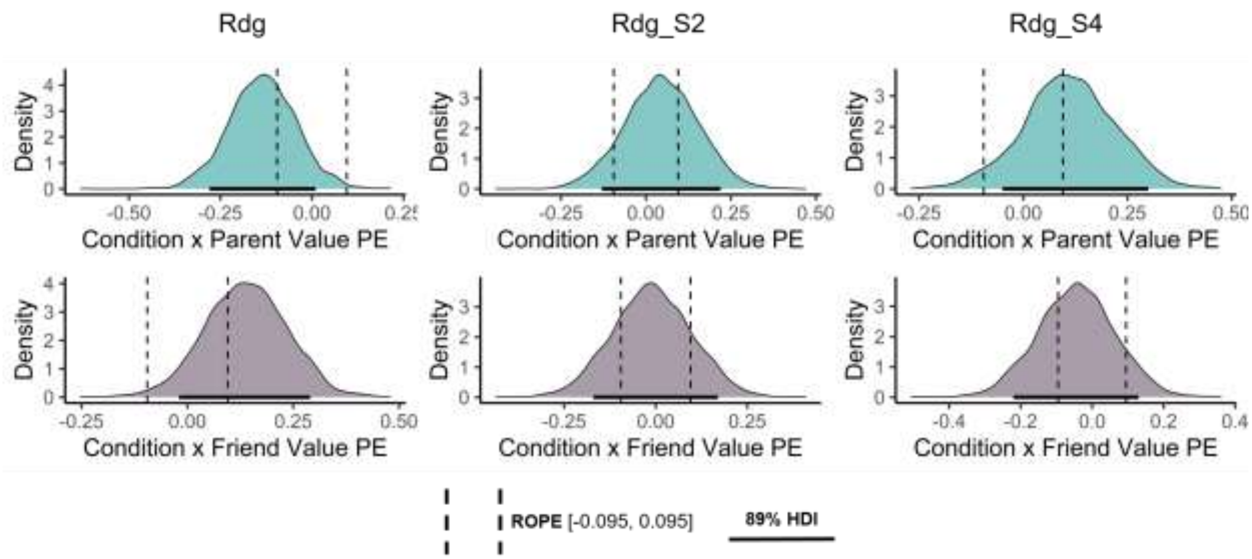
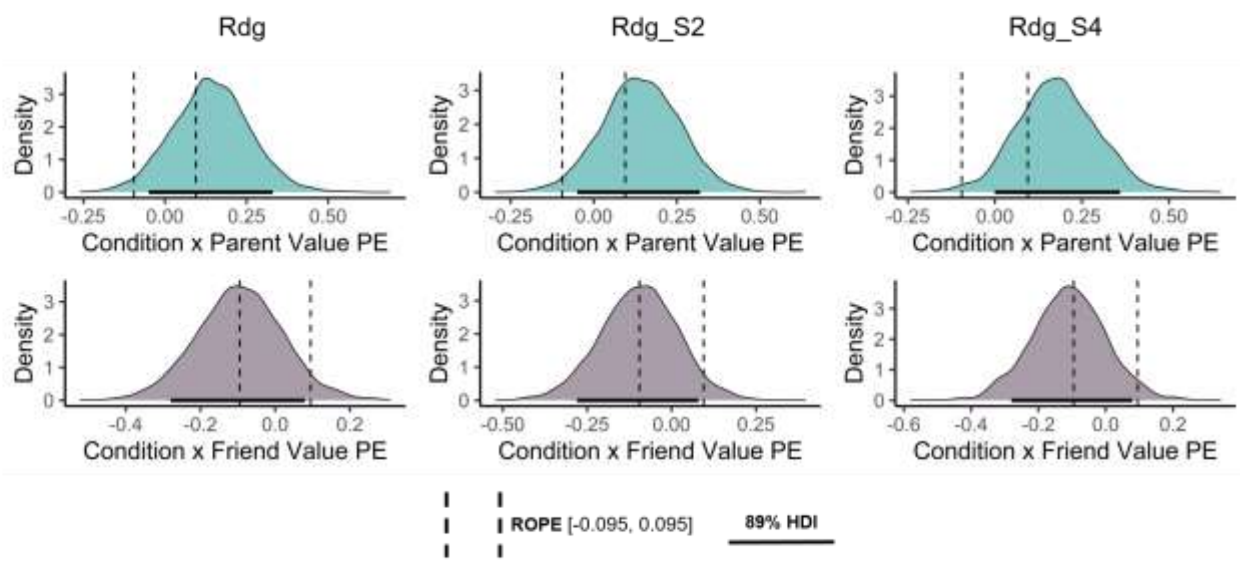


Figure A2.2. Posterior distribution plots for model interaction terms capturing the influence of value-based representations on social decision preferences (sample-specific neural signature, including only VS, mPFC).



Note. In both plots: 'Rdg' refers to the type of signature used (Rdg = sample-specific signature built using ridge regression). 'S' refers to the degree of smoothing in the custom signature (2 = 2mm, 4 = 4mm). 'PE' refers to pattern expression score. 'Condition x Parent/Friend' refers to the interaction term entered in the statistical model to assess the association between pattern expression scores and social decision preferences. 'ROPE' refers to Region of Practical Equivalence; 'HDI' refers to highest density credible intervals.

References

Abaied, J. L., & Emond, C. (2013). Parent psychological control and responses to interpersonal stress in emerging adulthood: Moderating effects of behavioral inhibition and behavioral activation. *Emerging Adulthood*, *1*(4), 258–270. https://doi.org/10.1177/2167696813485737

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2012). Human memory search as a random walk in a semantic network. *Neural Information Processing Systems,* 3050–3058.

Ahmed, S., Foulkes, L., Leung, J. T., Griffin, C., Sakhardande, A., Bennett, M., Dunning, D. L., Griffiths, K., Parker, J., Kuyken, W., Williams, G., Dalgleish, T., & Blakemore, S.-J. (2020). Susceptibility to prosocial and antisocial influence in adolescence. *Journal of Adolescence*, *84*, 56–68. https://doi.org/10.1016/j.adolescence.2020.07.012

Aknin, L. B., Barrington-Leigh, C. P., Dunn, E. W., Helliwell, J. F., Burns, J., Biswas-Diener, R., Kemeza, I., Nyende, P., Ashton-James, C. E., & Norton, M. I. (2013). Prosocial spending and well-being: Cross-cultural evidence for a psychological universal. *Journal of Personality and Social Psychology*, *104*(4), 635–652.

Albarello, F., Crocetti, E., & Rubini, M. (2018). I and Us: A Longitudinal Study on the Interplay of Personal and Social Identity in Adolescence. *Journal of Youth and Adolescence*, *47*, 689–702. https://doi.org/10.1007/s10964-017-0791-4

Almas, I., Cappelen, A. W., Sorensen, E. O., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*, 1176–1178.

Amodio, D. M. (2019). Social cognition 2.0: An interactive memory systems account. *Trends in Cognitive Sciences*, *23*(1), 21–33. https://doi.org/10.1016/j.tics.2018.10.002

Annas, J. (1993). *The morality of happiness*. Oxford University Press.

Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment: individual differences and their relationship to psychological well-being in adoles cence. *Journal of Youth & Adolescence*, *16*(5), 427–454. https://doi.org/10.1007/BF02202939

Armstrong-Carter, E., Guassi Moreira, J. F., Ivory, S. L., & Telzer, E. H. (2020). Supporting peers during adolescence is beneficial for most, but costly for some. *PsyArXiv*.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037//0003-066X.55.5.469

Arnett, J. J. (2014). *Adolescence and Emerging Adulthood*. Pearson.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology*, *60*(2), 241–253.

Arseniev-Koehler, A., Cochran, S. D., Mays, V. M., Chang, K.-W., & Foster, J. G. (2021). Integrating topic modeling and word embedding to characterize violent deaths. *ArXiv*.

Arseniev-Koehler, A., & Foster, J. G. (2020). Sociolinguistic properties of word embeddings. *SocArXiv*.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*(6), 359–372.

Avdeenko, A., & Gilligan, M. J. (2015). International interventions to build social capital: Evidence from a field experiment in Sudan. *American Political Science Review*, *109*(3), 427–449.

Ben-Ami Bartal, I., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science*, *334*(9), 1427–1430.

Ben-Ami Bartal, I., Rodgers, D. A., Bernardez Sarria, M. S., Decety, J., & Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. *ELife*, *3*, 1–16. https://doi.org/10.7554/eLife.01385

Benoit, W. L. (2011). Content analysis in political communication. In H. Bucy & L. Holbert (Eds.), *The Sourcebook for Political Communication Research* (pp. 268–279). Routledge.

Berkman, E. T., & Wilson, S. M. (2020). So useful as a good theory? *PsyArXiv*, 1–30.

Blakemore, S.-J. (2018). Avoiding social risk in adolescence. *Current Directions in Psychological Science*, *27*(2), 116–122. https://doi.org/10.1177/0963721417738144

Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Boer, D., & Boehnke, K. (2015). What are values? Where do they come from? A developmental Perspective. In T. Brosch & D. Sander (Eds.), *Handbook of Value: Perspectives from Economics, Neuroscience, Philosophy, Psychology, and Sociology* (pp. 129–151). Oxford University Press.

Botdorf, M., Rosenbaum, G. M., Patrianakos, J., Steinberg, L., & Chein, J. M. (2016). Adolescent risk-taking is predicted by individual differences in cognitive control over emotional, but not non-emotional, response conflict. *Cognition and Emotion*, 1–8. https://doi.org/10.1080/02699931.2016.1168285

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., … T., S. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*, 84–88. https://doi.org/10.1038/s41586-020-2314-9

Bouwmeester, S., Verkoeijen, P. P. J. L., Aczel, B., Barbosa, F., Bègue, L., Brañas-Garza, P., Chmura, T. G. H., Cornelissen, G., Døssing, F. S., Espín, A. M., Evans, A. M., Ferreira-Santos, F., Fiedler, S., Flegr, J., Ghaffari, M., Glöckner, A., Goeschl, T., Guo, L., Hauser, O. P., … Wollbrant, C. E. (2017). Registered Replication Report: Rand, Greene, and Nowak (2012). *Perspectives on Psychological Science*, *12*(3), 527–542. https://doi.org/10.1177/1745691617693624

Braams, B. R., & Crone, E. A. (2016). Peers and parents: A comparison between neural activation when winning for friends and mothers in adolescence. *Social Cognitive and Affective Neuroscience*, [EPub ahead of print].

Branje, S. J. T., Hale, W. W., Frijns, T., & Meeus, W. H. J. (2010). Longitudinal associations between perceived parent-child relationship quality and depressive symptoms in adolescence. *Journal of Abnormal Child Psychology*, *38*(6), 751–763. https://doi.org/10.1007/s10802-010-9401-6

Bray, S., Shimojo, S., & O'Doherty, J. P. (2010). Human medial orbitofrontal cortex is recruited during experience of imagined and real rewards. *Journal of Neurophysiology*, *103*(5), 2506–2512. https://doi.org/10.1152/jn.01030.2009

Brown, B. B., Lamborn, S. D., & Steinberg, L. (1993). Parenting practices and peer group affiliation in adolescence. *Child Development*, *64*(2), 467–482.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *186*, 183–186.

Camerer, C.F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, Colin F. (2015). Cognitive hierarchy process models of strategic thinking in games. *The Wiley Blackwell Handbook of Judgment and Decision Making*, *2*, 543–567.

Carlson, C. L. (2014). Seeking self-sufficiency: Why emerging adult college students receive and implement parental advice. *Emerging Adulthood*, *2*(4), 257–269. https://doi.org/10.1177/2167696814551785

Carlson, R. W., Aknin, L. B., & Liotti, M. (2016). When is giving an impulse? An ERP investigation of intuitive prosocial behavior. *Social Cognitive and Affective Neuroscience*, *11*(7), 1121–1129. https://doi.org/10.1093/scan/nsv077

Carlson, R. W., Marechal, M. A., Oud, B., Fehr, E., & Crockett, M. J. (2020). Motivated misremembering of selfish decisions. *Nature Communications*, 1–11. https://doi.org/10.1038/s41467-020-15602-4

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484.

Chang, L. J., Gianaros, P. J., Manuck, S. B., & Krishnan, A. (2015). A sensitive and specific neural signature for picture-induced negative affect. *PLOS Biology*, *13*(6), 1–28. https://doi.org/10.1371/journal.pbio.1002180

Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, *70*(3), 560–572.

Charest, I., Kievit, R. A., Schmitz, T. W., Deca, D., & Kriegeskorte, N. (2014). Unique semantic space in the brain of each beholder predicts perceived similarity. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14565–14570. https://doi.org/10.1073/pnas.1402594111

Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across child and adult language corpora of more than 65 million words. *Psychological Science*, *32*(2), 218–240. https://doi.org/10.1177/0956797620963619

Chavez, R. S., Heatherton, T. F., & Wagner, D. D. (2017). Neural population decoding reveals the intrinsic positivity of the self. *Cerebral Cortex*, *27*(11), 5222–5229. https://doi.org/10.1093/cercor/bhw302

Chen, G., Taylor, P. A., & Cox, R. W. (2017). Is the statistic value all we should care about in neuroimaging? *NeuroImage*, *147*, 952–959. https://doi.org/10.1101/064212

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics-Simulation and Computation*, *39*(4), 860–864. https://doi.org/10.1080/03610911003650383

Chen, M. K., Laskshminarayanan, V., & Santos, L. (2006). How basic are behavioral biases?

Evidence from capuchin monkey trading behavior. *Journal of Political Economy*, *114*(3), 517–537.

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

Contreras-Huertas, L. S., Lockwood, P. L., Bird, G., Apps, M., & Crockett, M. J. (2020). Prosocial behaviour is associated with transdiagnostic markers of affective sensitivity in multiple domains. *Emotion, in press*.

Conway, C. C., & Slavich, G. M. (2017). Behavior genetics of prosocial behavior. In *Compassion* (pp. 151–170). Routledge.

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*(48), 17320–17325. https://doi.org/10.1073/pnas.1424572112

Crockett, M. J., Siegel, J. Z., Kurth-Nelson, Z., Dayan, P., & Dolan, R. J. (2017). Moral transgressions corrupt neural representations of value. *Nature Neuroscience*, *May*, 1–10. https://doi.org/10.1038/nn.4557

Crone, E. A., & Fuligni, A. J. (2020). Self and Others in Adolescence. *Annual Review of Psychology*, *71*, 447–469.

Curtis, A. C. (2015). Defining adolescence. *Journal of Adolescent and Family Health*, *7*(2).

Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., & Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature*, *577*, 671–675. https://doi.org/10.1038/s41586-019-1924-6

Dal Monte, O., Chu, C. C. J., Fagan, N. A., & Chang, S. W. C. (2020). Specialized medial prefrontal-amygdala coordination in other-regarding decision preference. *Nature Neuroscience*, *23*, 565–574. https://doi.org/10.1038/s41593-020-0593-y

Damian, R. I., Spengler, M., Sutu, A., & Roberts, B. W. (2019). Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years. *Journal of Personality and Social Psychology*, *117*(3), 674.

DeFranza, D., Mishra, H., & Mishra, A. (2020). How language shapes prejudice against women: An examination across 45 world languages. *Journal of Personality and Social Psychology*, *119*(1), 7.

Delgado, M R, Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, *8*(11), 1611–1618. https://doi.org/10.1038/nn1575

Delgado, Mauricio R, Beer, J. S., Fellows, L. K., Huettel, S. A., Platt, M. L., Quirk, G. J., & Schiller, D. (2016). Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex. *Nature Neuroscience*, *19*(12), 1545–1552. https://doi.org/10.1038/nn.4438

DeStasio, K. L., Clithero, J. A., & Berkman, E. T. (2019). Neuroeconomics, health psychology, and the interdisciplinary study of preventative health behavior. *Social and Personality Psychology Compass*, *13*(10), 1–16. https://doi.org/10.1111/spc3.12500

Do, K. T., Guassi Moreira, J. F., & Telzer, E. H. (2017). But is helping you worth the risk? Defining prosocial risk taking in adolescence. *Developmental Cognitive Neuroscience*, *25*, 260–271. https://doi.org/10.1016/j.dcn.2016.11.008

Do, K. T., & Telzer, E. H. (2019). Developmental Cognitive Neuroscience Corticostriatal connectivity is associated with the reduction of intergroup bias and greater impartial giving in youth. *Developmental Cognitive Neuroscience*, *37*. https://doi.org/10.1016/j.dcn.2019.100628

Doré, B. P., Weber, J., & Ochsner, K. N. (2017). Neural predictors of decisions to cognitively control emotion. *Journal of Neuroscience*, *37*(10), 2580–2588. https://doi.org/10.1523/JNEUROSCI.2526-16.2016

Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid monte carlo. *Physics Letters B*, *195*(2), 216–222.

Dunham, Y. (2018). Mere Membership. *Trends in Cognitive Sciences*, *22*(9), 780–793.

DuPre, E., Hanke, M., & Poline, J.-B. (2020). Nature abhors a paywall: How open science can realize the potential of naturalistic stimuli. *NeuroImage*, *216*, 116330.

Efron, B., & Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association*, *70*(350), 311–319.

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. *Scientific American*, *236*(5), 119–127.

Eisenberg, N., Spinrad, T. L., & Knafo-Noam, A. (1998). *Prosocial Development*. John Wiley & Sons. https://doi.org/10.1002/9781118963418.childpsy315

Engelmann, J. B., Meyer, F. E. E., & Ruff, C. C. (2015). Anticipatory anxiety disrupts neural valuation during risky choice. *Journal of Neuroscience*, *35*(7), 3085–3099. https://doi.org/10.1523/JNEUROSCI.2880-14.2015

Etzel, J. A., Valchev, N., & Keysers, C. (2011). NeuroImage The impact of certain methodological choices on multivariate analysis of fMRI data with support vector machines. *NeuroImage*, *54*(2), 1159–1167. https://doi.org/10.1016/j.neuroimage.2010.08.050

Fanti, K. A., Henrich, C. C., Brookmeyer, K. A., & Kuperminc, G. P. (2008). Toward a Transactional Model of Parent-Adolescent Relationship Quality and Adolescent Psychological Adjustment. *Journal of Early Adolescence*, *28*(2), 252–276. https://doi.org/10.1177/0272431607312766

Fareri, D. S., Smith, D. V., & Delgado, M. R. (2020). The influence of relationship closeness on default-mode network connectivity during social interactions. *Social Cognitive and Affective Neuroscience*, *in press*, 1–16.

Feldmanhall, O., & Chang, L. J. (2018). Social Learning: Emotions Aid in Optimizing Goal-Directed Social Behavior. In R. Morris, A. Bornstein, & A. Shenhav (Eds.), *Goal-Directed Decision-Making: Computations and Circuits* (pp. 309–330). Academic Press. https://doi.org/10.1016/B978-0-12-812098-9.00014-0 ©

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

FeldmanHall, O., & Shenhav, A. (2018). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*, 426–435.

Feng, T., Zhao, W., & Donnay, G. F. (2013). The endowment effect can extend from self to mother: Evidence from an fMRI study. *Behavioural Brain Research*, *248*, 74–79.

Fernandez-Theoduloz, G., Paz, V., Nicolaisen-Sobesky, E., Perez, A., Buunk, A. P., Cabana, A., & Gradin, V. B. (2019). Social avoidance in depression: A study using a social decision-making task. *Journal of Abnormal Psychology*, *128*(3), 234–244.

Festinger, L. (1943). Studies in decision: II. An empirical test of a quantitative theory of decision. *Journal of Experimental Psychology*, *32*(5), 411–423.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 709–730. https://doi.org/10.1037/a0014983

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science*, *20*(4), 211–216. https://doi.org/10.1177/0963721411415790

Finn, E. S., Glerean, E., Khojandi, A. Y., Nielson, D., Molfese, P. J., Handwerker, D. A., & Bandettini, P. A. (2020). Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage*, *215*, 116828. https://doi.org/10.1016/j.neuroimage.2020.116828

Foulkes, L., Leung, J. T., Fuhrmann, D., Knoll, L. J., & Blakemore, S.-J. (2018). Age differences in the prosocial influence effect. *Developmental Science*, *21*, e12666. https://doi.org/10.1111/desc.12666

Freitas, A. L., Gollwitzer, P. M., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others self regulatory efforts. *Journal of Experimental Social Psychology*, *40*, 739–752.

Fromme, K., Corbin, W. R., & Kruse, M. I. (2008). Behavioral risks during the transition from high school to college. *Developmental Psychology*, *44*(5), 1497–1504. https://doi.org/10.1037/a0012614.Behavioral

Fu, X., Padilla-Walker, L. M., & Brown, M. N. (2017). Longitudinal relationships between adolescents’ self-esteem and prosocial behavior towards strangers, friends and family. *Journal of Adolescence*, *57*, 90–98.

Fuligni, A. J. (2018). The need to contribute during adolescence. *Perspectives on Psychological Science*, 1–13. https://doi.org/10.1177/1745691618805437

Fuligni, A. J., & Pedersen, S. (2002). Family obligation and the transition to young adulthood. *Developmental Psychology*, *38*(5), 856–868. https://doi.org/10.1037/0012-1649.38.5.856

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, *7*(3), 385–400.

Galvan, A., Hare, T. a, Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience*, *26*(25), 6885–6892. https://doi.org/10.1523/JNEUROSCI.1062-06.2006

Gangopadhyay, P., Chawla, M., Dal Monte, O., & Chang, S. W. C. (2021). Prefrontal-amygdala circuits in social decision-making. *Nature Neuroscience*, *24*, 5–18.

Gee, D. G., Gabard-Durnam, L., Telzer, E. H., Humphreys, K. L., Goff, B., Shapiro, M., Flannery, J., Lumian, D. S., Fareri, D. S., Caldera, C., & Tottenham, N. (2014). Maternal buffering of human amygdala-prefrontal circuitry during childhood but not during adolescence. *Psychological Science*, *25*(11), 2067–2078. https://doi.org/10.1177/0956797614550878

Gelman, A., Simpson, D., & Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy*, *19*(10), 555.

Gilbert, P. (2015). Affiliative and prosocial motives and emotions in mental health. *Dialogues in Clinical Neuroscience*, *17*(4), 381–389.

Gonzalez, B., & Chang, L. J. (2019). Computational models of mentalizing. *PsyArXiv*, 1–19.

Gould, R. L. (1972). The phases of adult life: A study in developmental psychology. *American Journal of Psychiatry*, *129*(5), 521–531. https://doi.org/10.1176/ajp.129.5.521

Grall, C., & Finn, E. S. (2021). The "naturalistic" fallcy: Leveraging the power of media to drive cognition. *PsyARXiv*.

Grecucci, A., Giorgetta, C., Van't Wout, M., Bonini, N., & Sanfey, A. G. (2013). Reappraising the ultimatum: An fMRI study of emotion regulation and decision making. *Cerebral Cortex*, *23*(2), 399–410. https://doi.org/10.1093/cercor/bhs028

Greenberger, E. (1984). Defining psychosocial maturity in adolescence. *Advances in Child Behavioral Analysis & Therapy*, *3*, 1–37.

Guassi Moreira, J. F., Tashjian, S. M., Galván, A., & Silvers, J. A. (2020). Is social decision making for close others consistent across domains and within individuals? *Journal of Experimental Psychology: General*, *149*(8), 1509–1526.

Guassi Moreira, J. F., & Telzer, E. H. (2018a). Family conflict is associated with longitudinal changes in striatal functional connectivity during adolescent risk taking under maternal influence. *Developmental Science*, *21*(5), 1–10. https://doi.org/10.1111/desc.12632

Guassi Moreira, J. F., & Telzer, E. H. (2018b). Family conflict shapes how adolescents take risks when their family is affected. *Developmental Science*, *21*(4), 1–9. https://doi.org/10.1111/desc.12611

Guassi Moreira, J. F., & Telzer, E. H. (2018c). Longitudinal increases in family relationship quality and sensation seeking interact to increase risk raking in emerging adults. *Emerging Adulthood*, *6*(1), 66–71.

Guthrie, T., Benadjaoud, Y. Y., & Chavez, R. S. (2021). Social relationship strength modulates the similarity of brain-to-brain representations of group members. *PsyArXiv*.

Haber, S. N., & Knutson, B. (2009). The Reward Circuit: Linking Primate Anatomy and Human Imaging. *Neuropsychopharmacology*, *35*(10), 4–26. https://doi.org/10.1038/npp.2009.129

Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards : dissociable neural correlates and effects on choice. *Nature Neuroscience*, *18*(9), 1233–1235. https://doi.org/10.1038/nn.4080

Hackel, L. M., & Zaki, J. (2018). Propagation of economic inequality through reciprocity and reputation. *Psychological Science*, *29*(4), 604–613. https://doi.org/10.1177/0956797617741720

Hackel, L. M., Zaki, J., & Van Bavel, J. J. (2017). Social identity shapes social valuation: evidence from prosocial behavior and vicarious reward. *Social Cognitive and Affective Neuroscience*, *12*(8), 1219–1228. https://doi.org/10.1093/scan/nsx045

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G. A., Bruyneel, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T. T. L., Crowell, A., De Ridder, D. T. D., Dewitte, S., … Zwienenberg, M. (2016). A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

Hall, G. C. N., Berkman, E. T., Zane, N. W., Leong, Frederick, T. L., Hwang, W.-C., Nezu, A. M., Nezu, C. M., Hong, J. J., Chu, J. P., & Huang, E. R. (2020). Reducing Mental Health Disparities by Increasing the Personal Relevance of Interventions. *American Psychologist*, 1–39.

Harle, K. M., Allen, J. J. B., & Sanfey, A. G. (2010). The impact of depresssion on social economic decision making. *Journal of Abnormal Psychology*, *119*(2), 440–446.

Harris, A., Clithero, J. A., & Hutcherson, C. A. (2018). Accounting for taste: A multi-attribute neurocomputational model explains the neural dynamics of choices for self and others. *The Journal of Neuroscience*, *38*(37), 7952–7968. https://doi.org/10.1523/JNEUROSCI.3327-17.2018

Hassabis, D., Spreng, R. N., Rusu, A. A., Robbins, C. A., Mar, R. A., & Schacter, D. L. (2014). Imagine all the people: How the brain creates and uses personality models to predict behavior. *Cerebral Cortex*, *24*(8), 1979–1987. https://doi.org/10.1093/cercor/bht042

Helmer, M., Warrington, S., Mohammadi-Nejad, A.-R., Ji, J. L., Howell, A., Rosand, B., Anticevic, A., Sotiropoulos, S. N., & Murray, J. D. (2020). On stability of canonical correlation analysis and partial least squares with application to brain-behavior associations. *BioRxiv*, 1–43.

Henrich, J. (2002). Decision-making, cultural transmission and adaptation in economic anthropology. In J. Ensminger (Ed.), *Theory in Economic Anthropology* (pp. 251–295). AltaMira Press.

Higgins, E. T. (2015). What is value? Where does it come from? A psychological perspective. In T. Brosch & D. Sander (Eds.), *Handbook of Value: Perspectives from Economics, Neuroscience, Philosophy, Psychology, and Sociology* (pp. 43–63). Oxford University Press.

Hine, J. W., & Usynin, A. (2005). MSET performance optimization through regularization. *Nuclear Engineering and Technology*, *37*(2), 177–184.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research, 15*(1), 1593–1623.

Hong, Y., Yoo, Y., Han, J., Wager, T. D., & Woo, C.-W. (2019). False-positive neuroimaging: Undisclosed flexibility in testing spatial hypotheses allows presenting anything as a replicated finding. *NeuroImage*, *195*, 384–395.

Hopmeyer, A., & Medovoy, T. (2017). Emerging adults' self-identified peer crowd affiliations and college adjustment. *Emerging Adulthood*, *5*(2), 143–148. https://doi.org/10.1007/s11218-017-9390-1

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences*, *32*(3), 401–414. https://doi.org/10.1016/S0191-8869(01)00032-0

Huck, S., Lünser, G. K., & Tyran, J. (2012). Competition fosters trust. *Games and Economic*

Behavior, *76*(1), 195–209. https://doi.org/10.1016/j.geb.2012.06.010

Hudson, N. W., Fraley, R. C., Chopnik, W. J., & Heffernan, M. E. (2015). Not all attachment relationships develop alike: Normative cross-sectional age trajectories in attachment to romantic partners, best friends, and parents. *Journal of Research in Personality*, *59*, 44–55.

Hull, C. L. (1931). Goal attraction and directing ideas conceied as habit phenomena. *Psychological Review*, *38*(6), 487–506.

Huth, A. G., Heer, W. A. De, Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637.Natural

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210–1224. https://doi.org/10.1016/j.neuron.2012.10.014

Israel, S., Hasenfratz, L., & Knafo-Noam, A. (2015). The genetics of morality and prosociality. *Current Opinion in Psychology*, *6*, 55–59.

Jackson, J. C., Watts, J., List, J.-M., Puryear, C., Drabble, R., & Lindquist, K. A. (2021). From text to thought: How analyzing language can advance psychological science. *Perspectives on Psychological Science*, 1–22.

James, W., & Stein, C. (1992). Estimation with quadratic loss. In *Breakthroughs in Statistics* (pp. 443–460). Springer.

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*(5), 865–889. https://doi.org/10.1016/j.joep.2011.05.007

Jolly, E., & Chang, L. J. (2019). The flatland fallacy: Moving beyond low dimensional thinking. *Topics in Cognitive Science*, *11*(2), 433–454.

Juckett, D. (2012). A method for determining the number of documents needed for a gold standard corpus. *Journal of Biomedical Informatics*, *45*(3), 460–470. https://doi.org/10.1016/j.jbi.2011.12.010

Kahn, F., Fraley, R. C., Young, J. F., & Hankin, B. L. (2019). Developmental traectories of attachment and depressive symptoms in children and adolescents. *Attachment & Human Development*, 1–17.

Kang, M. J., Rangel, A., Camus, M., & Camerer, C. F. (2011). Hypothetical and real choice differentially activate common valuation areas. *Journal of Neuroscience*, *31*(2), 461–468. https://doi.org/10.1523/JNEUROSCI.1583-10.2011

Kaniušonytė, G., & Žukauskienė, R. (2018). Relationships with parents, identity styles, and positive youth development during the transition from adolescence to emerging adulthood. *Emerging Adulthood*, *6*(1), 42–52. https://doi.org/10.1177/2167696817690978

Kappes, A., Nussberger, A.-M., Faber, N. S., Kahane, G., Savulescu, J., & Crockett, M. J. (2018). Uncertainty about the impact of social decisions increases prosocial behaviour. *Nature Human Behaviour*, *2*, 573–580. https://doi.org/10.1038/s41562-018-0372-x

Kerr, B., Godfrey-Smith, P., & Feldman, M. (2004). What is altruism? *Trends in Ecology & Evolution*, *19*(3), 135–140.

Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, *14*(3), 608–638.

Knafo, A., Steinberg, T., & Goldner, I. (2011). Children's low affective perspective-taking ability is associated with low self-initiated pro-sociality. *Emotion*, *11*(1), 194.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport*, *12*(17), 3683–3687. https://doi.org/10.1097/00001756-200112040-00016

Krajbich, I., Adolphs, R., Tranel, D., Denburg, N. L., & Camerer, C. F. (2009). Economic cames quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *The Journal of Neuroscience*, *29*(7), 2188–2192. https://doi.org/10.1523/JNEUROSCI.5086-08.2009

Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, *21*, 1148–1160.

Krishnan, A., Woo, C.-W., Chang, L. J., Ruzic, L., Gu, X., Lopez-Sola, M., Jackson, P. L., Pujol, J., Fan, J., & Wager, T. D. (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *ELife*, *5*, e15166. https://doi.org/10.7554/eLife.15166

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*(3), 299–312.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t-test. *Journal of Experimental Psychology: General*, *142*(2), 573–603.

Kurdi, B., Mann, T. C., Charlesworth, T. E. S., & Banaji, M. R. (2019). The relationship between implicit intergroup attitudes and beliefs. *Proceedings of the National Academy of Sciences*, *116*(13), 5862–5871. https://doi.org/10.1073/pnas.1820240116

Kutlu, M. G., Zachry, J. E., Melugin, P. R., Cajigas, S. A., Chevee, M. F., Kelley, S. J., Kutlu, B., Tian, L., Siciliano, C. A., & Calipari, E. S. (2021). Dopamine release in the nucleus accumbens core signals perceived saliency. *Current Biology*, 1–14. https://doi.org/10.1016/j.cub.2021.08.052

Kwak, Y., & Huettel, S. A. (2016). Prosocial reward learning in children and adolescents. *Frontiers in Psychology*, *7*, 1–10. https://doi.org/10.3389/fpsyg.2016.01539

Kwak, Y., Pearson, J., & Huettel, S. A. (2014). Differential reward learning for self and others predicts self-reported altruism. *PLoS ONE*, *9*(9), 1–9. https://doi.org/10.1371/journal.pone.0107621

Lamba, A., Frank, M. J., & FeldmanHall, O. (2020). Anxiety impedes adaptive social learning under uncertainty. *Psychological Science*. https://doi.org/10.1177/0956797620910993

Levin, I. P., & Hart, S. S. (2003). Risk preferences in young children: Early evidence of individual differences in reaction to potential gains and losses. *Journal of Behavioral Decision Making*, *16*(5), 397–413.

Levinson, D. J. (1986). A conception of adult development. *American Psychologist*, *41*(1), 3–13. https://doi.org/10.1037/0003-066X.41.1.3

Lewin, K. (1938). *Contributions to psychological theory: The conceptual representation and the measurement of psychological forces*. Duke University Press. https://doi.org/https://doi.org/10.1037/13613-000

Lin, X., & Boutros, P. C. (2020). Optimization and expansion of non-negative matrix factorization. *BMC Bioinformatics*, *21*(7), 1–10.

Lindström, B., Bellander, M., Schultner, D. T., Chang, A., Tobler, P. N., & Amodio, D. M. (2021). media engagement. *Nature Communications*, 1–10.

Locey, M. L., Jones, B. A., & Rachlin, H. (2011). Real and hypothetical rewards. *Judgment and Decision Making*, *6*(6), 552–264. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3348706/%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/3348706

Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences*, *113*(35), 9763–9768. https://doi.org/10.1073/pnas.1603198113

Lockwood, P. L., Hamonet, M., Zhang, S. H., Ratnavel, A., Salmony, F. U., & Husain, M. (2017). Prosocial apathy for helping others when effort is required. *Nature Human Behaviour*, *1*(0131), 1–10. https://doi.org/10.1038/s41562-017-0131

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., & Pessiglione, M. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nature Neuroscience*, *23*, 664–675. https://doi.org/10.1038/s41593-020-0615-9

Lucas-Thompson, R. G. (2014). Relationship quality with parents, stressful life events, and cortisol production in emerging adulthood. *Emerging Adulthood*, *2*(2), 92–104. https://doi.org/10.1177/2167696813503313

Makowski, D., Ben-Shachar, M. S., & Ludecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Marek, A. S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Feczko, E., Miranda-Domínguez, Ó., Graham, A. M., Earl, E., Perrone, A. J., Cordova, M., Doyle, O., Moore, L. A., Conan, G., Uriarte, J., … Dosenbach, N. U. F. (2020). Towards reproducible brain-wide association studies. *BioRxiv*, 1–40.

Marek, S., & Dosenbach, N. U. F. (2018). The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues in Clinical Neuroscience*, *20*, 133–140. www.dialogues-cns.org

Mathews, H. F. (1987). Predicting decision outcomes: Have we put the cart before the horse in anthropological studies of decision making? *Human Organization*, *46*(1), 54–61.

Mattanah, J. F., Lopez, F. G., & Govern, J. M. (2011). The contributions of parental attachment bonds to college student development and adjustment: A meta-analytic review. *Journal of Counseling Psychology*, *58*(4), 565–596. https://doi.org/10.1037/a0024635

Mayr, U., & Freund, A. M. (2020). Do we become more prosocial as we age, and if so, why? *Current Directions in Psychological Science*, 1–7. https://doi.org/10.1177/0963721420910811

McCarthy, R. J., Skowronski, J. J., Verschuere, B., Meijer, E. H., Jim, A., Hoogesteyn, K., Orthey, R., Acar, O. A., Aczel, B., Bakos, B. E., Barbosa, F., Baskin, E., Bègue, L., Ben-Shakhar, G., Birt, A. R., Blatz, L., Charman, S. D., Claesen, A., Clay, S. L., … Yıldız, E.

(2018). Registered Replication Report on Srull and Wyer (1979). *Advances in Methods and Practices in Psychological Science*, *1*(3), 321–336. https://doi.org/10.1177/2515245918777487

McElreath, R. (2018). *Statistical rethinking: A bayesian course with examples in R and Stan* (1st ed.). Chapman and Hall/CRC. https://www.taylorfrancis.com/books/mono/10.1201/9781315372495/statistical-rethinking-richard-mcelreath

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology1*, *46*, 806–834. https://doi.org/10.1016/j.appsy.2004.02.002

Meyer, D. J. C., & Anderson, H. C. (2000). Preadolescents and apparel purchasing: Conformity to parents and peers in the consumer socialization process. *Journal of Social Behavior and Personality*, *15*(2), 243–258.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *ArXiv*.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*(4), 371–378.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40–48.

Misaki, M., Luh, W.-M., & Bandettini, P. A. (2014). The effect of spatial smoothing on fMRI decoding of columnar-level organization with linear support vector machine. *Journal of*

Neuroscience Methods, *212*(2), 355–361. https://doi.org/10.1016/j.jneumeth.2012.11.004.The

Miyapuram, K. P., Tobler, P. N., Gregorios-Pippas, L., & Schultz, W. (2012). BOLD responses in reward regions to hypothetical and imaginary monetary rewards. *NeuroImage*, *59*(2), 1692–1699. https://doi.org/10.1016/j.neuroimage.2011.09.029

Morelli, S. A., Chang, Y., Carlson, R. W., Kullar, M., & Zaki, J. (2018). Neural detection of socially valued community members. *Proceedings of the National Academy of Sciences*, *115*(32), 8149–8154. https://doi.org/10.1073/pnas.1712811115

Mumford, J. A., Davis, T., & Poldrack, R. A. (2014). The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *NeuroImage*, *103*, 130–138. https://doi.org/10.1016/j.neuroimage.2014.09.026

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*, *59*(3), 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076

Neal, R. M. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, *2*(11).

Needham, B. L. (2008). Reciprocal relationships between symptoms of depression and parental support during the transition from adolescence to young adulthood. *Journal of Youth and Adolescence*, *37*(8), 893–905. https://doi.org/10.1007/s10964-007-9181-7

Niv, Y., & Chan, S. (2011). On the value of information and other rewards. *Nature Neuroscience*, *14*(9), 1095–1097. https://doi.org/10.1038/nn.2918

Padilla-Walker, L. M., Dyer, W. J., Yorgason, J. B., Fraser, A. M., & Coyne, S. M. (2013). Adolescents' prosocial behavior toward family, friends, and strangers: A person-centered approach. *Journal of Research on Adolescence*, *25*(1), 135–150.

Padilla-Walker, L. M., & Fraser, A. M. (2014). How much is it going to cost me? Bidirectional relations between adolescents' moral personality and prosocial behavior. *Journal of Adolescence*, *37*(7), 993–1001.

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2017). Spontaneous neural encoding of social network position. *Nature Human Behaviour*, *1*(5), 1–7. https://doi.org/10.1038/s41562-017-0072

Pärnamets, P., Shuster, A., Reinero, D. A., & Van Bavel, J. J. (2019). A value-based framework for understanding cooperation. *PsyArXiv*, 1–15.

Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science*, *23*(6), 421–426. https://doi.org/10.1177/0963721414547737

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pfeifer, J. H., & Allen, N. B. (2012). Arrested Development? Reconsidering dual-systems models of brain function in adolescence and disorders. *Trends in Cognitive Sciences*, *16*(6), 322–329. https://doi.org/10.1038/jid.2014.371

Pfeifer, J. H., & Berkman, E. T. (2018). The Development of Self and Identity in Adolescence: Neural Evidence and Implications for a Value-Based Choice Perspective on Motivated Behavior. *Child Development Perspectives*, *12*(3), 158–164.

https://doi.org/10.1111/cdep.12279

Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D. (2009). Neural correlates of direct and reflected self-appraisals in adolescents and adults: When social perspective-taking informs self-perception. *Child Development*, *80*(4), 1016–1038. https://doi.org/10.1038/jid.2014.371

Phillips, D. L., & Clancy, K. J. (1972). Some effects of “social desirability” in survey studies. *American Journal of Sociology*, *77*(5).

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data. *Trends in Cognitive Sciences*, *10*(2), 59–63.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M., Nichols, T. E., Poline, J.-B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18,* 115–126. https://doi.org/http://dx.doi.org/10.1101/059188

Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M., Gorgolewski, K. J., Luci, J., Joo, S. J., Boyd, R. L., Hunicke-smith, S., Simpson, Z. B., Caven, T., Sochat, V., Shine, J. M., Gordon, E., Snyder, A. Z., Adeyemo, B., Petersen, S. E., … Mumford, J. A. (2015). Long-term neural and physiological phenotyping of a single human. *Nature Communications*, *6*, 8885. https://doi.org/10.1038/ncomms9885

Poulin, M. J., & Haase, C. M. (2015). Growing to trust: Evidence that trust increases and sustains well-being across the life span. *Social Psychological and Personality Science*, *6*(6), 614–621. https://doi.org/10.1177/1948550615574301

Powers, K. E., Yaffe, G., Hartley, C. A., Davidow, J. Y., Kober, H., & Somerville, L. H. (2018). Consequences for peers differentially bias computations about risk across development. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000389

Pronin, E., Olivola, C. Y., & Kennedy, K. A. (2008). Doing unto future selves as you would do unto others: Psychological distance and decision making. *Personality and Social Psychology Bulletin*, *34*(2), 224–236. https://doi.org/10.1177/0146167207310023

Rhoads, S. A., Cutler, J., & Marsh, A. A. (2021). A feature-based network analysis and fMRI meta-analysis reveal three distinct types of prosocial decisions. *BioRxiv*.

Richards, J. M., Plate, R. C., & Ernst, M. (2013). A systematic review of fMRI reward paradigms used in studies of adolescents vs. adults: The impact of task design and implications for understanding neurodevelopment. *Neuroscience and Biobehavioral Reviews*, *37*(5), 976–991. https://doi.org/10.1016/j.neubiorev.2013.03.004

Riedijk, L., & Harakeh, Z. (2017). Imitating the risky decision-making of peers. *Emerging Adulthood*, 1–11. https://doi.org/10.1177/2167696817722918

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, *62*, 23–48. https://doi.org/10.1080/1068316X.2017.1414817

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*(3), 582–593.

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality*, *69*(4), 617–640.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–151.

Roisman, G. I., Masten, A. S., Coatsworth, J. D., & Tellegen, A. (2004). Salient and emerging developmental tasks in the transition to adulthood. *Child Development*, *75*(1), 123–133.

Rotter, J. B. (1960). Some implications of a social learning theory for the prediction of goal directed behavior from testing procedures. *Psychological Review*, *67*(5), 301–316.

Ruhl, H., Dolan, E. A., & Buhrmester, D. (2015). Adolescent attachment trajectories with mothers and fathers: The importance of parent-child relationship experiences and gender. *Journal of Research on Adolescence*, *25*(3), 427–442.

Salehi, M., Greene, A. S., Karbasi, A., Shen, X., Scheinost, D., & Constable, T. (2020). There is no single functional atlas for even a single individual: Functional parcel definitions change with task. *NeuroImage*, *208*, 116366.

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet: Child & Adolescent Health*, *2*(3), 223–228.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347–367.

Schnakenberg, K. E., & Turner, I. R. (2019). Signaling with reform: How the threat of corruption prevents informed policy-making. *American Political Science Review*, *113*(3), 762–777.

Schnakenberg, K. E., & Turner, I. R. (2021). Helping friends or influencing foes: Electoral and policy effects of campaign finance contributions. *American Journal of Political Science*, *65*(1), 88–100.

Scholl, J., Kolling, N., Nelissen, N., Wittmann, M. K., Harmer, C. J., & Rushworth, M. F. S. (2015). The Good, the bad, and the irrelevant: Neural mechanisms of learning real and hypothetical rewards and effort. *Journal of Neuroscience*, *35*(32), 11233–11251. https://doi.org/10.1523/JNEUROSCI.0396-15.2015

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, *4*, 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Seaman, K. L., Gorlick, M. A., Vekaria, K. M., Hsu, M., Zald, D. H., & Samanez-Larkin, G. R. (2016). Adult age differences in decision-making across domains: Increased discounting of social and health-related rewards. *Psychology and Aging*, *31*(7), 737–746.

Shah, J. (2003). Automatic for people: How representations of significant others implicitly affect goal pursuit. *Journal of Personality and Social Psychology*, *84*, 661–681.

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*, 103–117. https://doi.org/10.1016/j.dcn.2015.12.010

Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., & Petersen, S. E. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Human Brain Mapping*, *35*(5), 1981–1996. https://doi.org/10.1002/hbm.22307.Statistical

Sierksma, J., & Shutts, K. (2020). When helping hurts: Children think groups that receive help are less smart. *Child Development*, *91*(3), 715–723.

Sierksma, J., & Shutts, K. (2021). Competence-based helping: Children's consideration of need when providing others with help. *Journal of Experimental Child Psychology*, *210*, 105206.

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahnik, S., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., … Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Silva, K., Chein, J., & Steinberg, L. (2016). Adolescents in peer groups make more prudent decisions when a slightly older adult is present. *Psychological Science, 27*(3), 322–330. https://doi.org/10.1177/0956797615620379

Silva, K., Chein, J., & Steinberg, L. (2020). The influence of romantic partners on male. *Journal of Social and Personal Relationships*, 1–11. https://doi.org/10.1177/0265407519899712

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, *12*(6), 1123–1128. https://doi.org/10.1177/1745691617708630

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *32*(1), 1–28.

Simpson, J. A. (2007). Psychological foundations of trust. *Current Directions in Psychological Science*, *16*(5), 264–268.

Sims, C. R. (2018). Efficient coding explains the universal law of generalization in human perception. *Science*, *656*, 652–656.

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience, 8*, 341–350. https://doi.org/10.1093/scan/nss002

Sokol-Hessner, P., Hutcherson, C., Hare, T., & Rangel, A. (2012). Decision value computation in DLPFC and VMPFC adjusts to the available decision time. *European Journal of Neuroscience*, *35*, 1065–1074. https://doi.org/10.1111/j.1460-9568.2012.08076.x

Sokol-Hessner, P., & Rutledge, R. (2019). The psychological and neural basis of loss aversion. *Current Directions in Psychological Science*, *28*(1), 20–27.

Spear, L. P. (2000a). Neurobehavioral Changes in Adolescence. *Current Directions in Psychological Science*, *9*(4), 111–114.

Spear, L. P. (2000b). *The adolescent brain and age-related behavioral manifestations* (Vol. 24).

Stamatis, C. A., Engelmann, J. B., Ziegler, C., Domschke, K., Hasler, G., & Timpano, K. R. (2020). A neuroeconomic investigation of 5-HTT/5-HT1A gene variation, social anxiety, and risk-taking behavior. *Anxiety, Stress, & Coping*, *33*(2), 176–192.

Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., Chang, L., Chaudhary, N., Giunta, L. Di, Dodge, K. A., Fanti, K. A., Lansford, J. E., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., … Takash, H. M. S. (2017). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science, [EPub Ahead of Print]*, 1–13. https://doi.org/10.1111/desc.12532

Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology*, *52*, 83–110. https://doi.org/10.1146/annurev.psych.52.1.83

Stiglic, G., Povalej Brzan, P., Fijacko, N., Wang, F., Delibasic, B., Kalousis, A., & Obradovic, Z. (2015). Comprehensible predictive modeling using regularized logistic regression and comorbidity based features. *PLoS One*, *10*(12), e0144439.

Strang, S., Hoeber, C., Uhl, O., Koletzko, B., Münte, T. F., Lehnert, H., Dolan, R. J., Schmid, S. M., & Park, S. Q. (2017). Impact of nutrition on social decision making. *Proceedings of the National Academy of Sciences*, *114*(25), 6510–6514. https://doi.org/10.1073/pnas.1620245114

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*, 159–171.

Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social mind. *Trends in Cognitive Sciences*, *22*(3), 201–212.

Taylor, M. J., Arsalidou, M., Bayless, S. J., Morris, D., Evans, J. W., & Barbeau, E. J. (2009). Neural correlates of personally familiar faces: Parents, partner and own faces. *Human Brain Mapping*, *30*(7), 2008–2020. https://doi.org/10.1002/hbm.20646

Tebes, J. K. (2000). External validity and scientific psychology. *American Psychologist*, *55*(12), 1508–1509.

Telzer, E. H., Gonzales, N., & Fuligni, A. J. (2014). Family obligation values and family assistance behaviors: Protective and risk factors for Mexican-American adolescents' substance use. *Journal of Youth and Adolescence*, *43*(2), 270–283. https://doi.org/10.1007/s10964-013-9941-5

Telzer, E. H., Ichien, N. T., & Qu, Y. (2015). Mothers know best: Redirecting adolescent reward sensitivity toward safe behavior during risk taking. *Social Cognitive and Affective Neuroscience*, *10*(10), 1383–1391. https://doi.org/10.1093/scan/nsv026

Telzer, E. H., Masten, C. L., Berkman, E. T., Lieberman, M. D., & Fuligni, A. J. (2010). Gaining while giving: an fMRI study of the rewards of family assistance among white and Latino youth. *Social Neuroscience*, *5*(5–6), 508–518. https://doi.org/10.1080/17470911003687913

Telzer, E. H., Masten, C. L., Berkman, E. T., Lieberman, M. D., & Fuligni, A. J. (2011). Neural regions associated with self control and mentalizing are recruited during prosocial behaviors towards the family. *NeuroImage*, *58*(1), 242–249. https://doi.org/10.1016/j.neuroimage.2011.06.013

Theisen, J. C., Fraley, R. C., Hankin, B. L., Young, J. F., & Chopnik, W. J. (2018). How do attachment styles change from childhood through adolescence? Findings from an accelerated longitudinal cohort study. *Journal of Research in Personality*, *74*, 141–146.

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin*, *146*(1), 30–90.

Ting, C.-C., Palminteri, S., Lebreton, M., & Engelmann, J. B. (2021). The elusive effects of incidental anxiety on reinforcement-learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2009). Risk-dependent reward value signal in human prefrontal cortex. *Proceedings of the National Academy of Sciences*, *106*(17), 7185–7190.

Torrance, E. P. (1957). Group decision-making and disagreement. *Social Forces*, *35*(4), 314–318.

Tottenham, N. (2020). Early adversity and the neotenous human brain. *Biological Psychiatry*, *87*(4), 350–358.

Tsai, K. M., Telzer, E. H., & Fuligni, A. J. (2013). Continuity and discontinuity in perceptions of family relationships from adolescence to young adulthood. *Child Development*, *84*(2), 471–484. https://doi.org/10.1111/j.1467-8624.2012.01858.x

Tversky, A., & Hutchinson, J. W. (1986). Nearest neighbor analysis of psychological spaces. *Psychological Review*, *93*(1), 3–22.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Uy, J., & Galván, A. (2017). Acute stress increases risky decisions and dampens prefrontal activation among adolescent boys. *NeuroImage*, *146*, 679–689. https://doi.org/10.1016/j.neuroimage.2016.08.067

Vaish, A., Carpenter, M., & Tomasello, M. (2016). The early emergence of guilt-motivated prosocial behavior. *Child Development*, *87*(6), 1772–1782.

https://doi.org/10.1111/cdev.12628

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias. *Psychological Science*, *19*(11), 1131–1139.

van Duijvenvoorde, A. C. K., Huizenga, H. M., Somerville, L. H., Delgado, M. R., Powers, A., Weeda, W. D., Casey, B. J., Weber, E. U., & Figner, B. (2015). Neural correlates of expected risks and returns in risky choice across development. *The Journal of Neuroscience*, *35*(4), 1549–1560. https://doi.org/10.1523/JNEUROSCI.1924-14.2015

Van Hoorn, J., Crone, E. A., & Van Leijenhorst, L. (2017). Hanging out with the right crowd: Peer influence on risk-taking behavior in adolescence. *Journal of Research on Adolescence*, *27*(1), 189–200. https://doi.org/10.1111/jora.12265

van Hoorn, J., van Dijk, E., Meuwese, R., Rieffe, C., & Crone, E. A. (2016). Peer influence on prosocial behavior in adolescence. *Journal of Research on Adolescence*, *26*(1), 90–100. https://doi.org/10.1111/jora.12173

Verschuere, B., Meijer, E. H., Jim, A., Hoogesteyn, K., Orthey, R., McCarthy, R. J., Skowronski, J. J., Acar, O. A., Aczel, B., Bakos, B. E., Barbosa, F., Baskin, E., Bègue, L., Ben-Shakhar, G., Birt, A. R., Blatz, L., Charman, S. D., Claesen, A., Clay, S. L., … Yıldız, E. (2018). Registered Replication Report on Mazar, Amir, and Ariely (2008). *Advances in Methods and Practices in Psychological Science*, 251524591878103. https://doi.org/10.1177/2515245918781032

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*(11), 546–554. https://doi.org/10.1016/j.tics.2011.09.008

Wager, T. D., Atlas, L. Y., Lindquist, M. A., Roy, M., Woo, C.-W., & Kross, E. (2013). An fMRI-based neurologic signature of physical pain. *New England Journal of Medicine*, *368*(15), 1388–1397.

Wake, S. J., & Izuma, K. (2017). A common neural code for social and monetary rewards in the human striatum. *Social Cognitive and Affective Neuroscience*, *12*(10), 1558–1564. https://doi.org/10.1093/scan/nsx092

Wang, S.-M., Cheng, Y.-H., Lee, C.-F., & Chuang, S.-C. (2019). Endowment effect: Trading for oneself versus trading for others. *Psychological Reports*, *122*(6), 2298–2319. https://doi.org/10.1177/0033294118802555

Wang, Y., Collins, J. A., Koski, J., Nugiel, T., Metoki, A., & Olson, I. R. (2017). Dynamic neural architecture for social knowledge retrieval. *Proceedings of the National Academy of Sciences*, *114*(16), E3305–E3314. https://doi.org/10.1073/pnas.1621234114

Weinstein, N., & Ryan, R. M. (2010). When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *Journal of Personality and Social Psychology*, *98*(2), 222–244.

Welborn, B. L., Lieberman, M. D., Goldenberg, D., Fuligni, A. J., Galván, A., & Telzer, E. H. (2015). Neural mechanisms of social influence in adolescence. *Social Cognitive and Affective Neuroscience*, *11*(1), 100–109. https://doi.org/10.1093/scan/nsv095

White, H. R., McMorris, B. J., Catalano, R. F., Fleming, C. B., Haggerty, K. P., & Abbott, R. D. (2006). Increases in alcohol and marijuana use during the transition out of high school into emerging adulthood: The effects of leaving home, going to college, and high school protective factors. *Journal of Studies on Alcohol*, *67*(6), 810–822.

https://doi.org/10.15288/jsa.2006.67.810

Xiao, J., Chen, L., Yu, Y., Zhang, X., & Chen, J. (2018). A phylogeny-regularized sparse regression model for predictive modeling of microbial community data. *Frontiers in Microbiology*, *9*, 3112.

Yao, B., & Yang, H. (2016). Physics-driven spatiotemporal regularization for high-dimensional predictive modeling: A novel approach to solve the inverse ECG problem. *Scientific Reports*, *6*(1), 1–13.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–670.

Yeager, D. S., Dahl, R. E., & Dweck, C. S. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, *13*(1), 101–122. https://doi.org/10.1177/1745691617722620

Yoder, K. J., & Decety, J. (2018). The neuroscience of morality and social decision-making. *Psychology, Crime and Law*, *24*(3), 279–295. https://doi.org/10.1080/1068316X.2017.1414817

Zaki, J., & Mitchell, J. P. (2013). Intuitive Prosociality. *Current Directions in Psychological Science*, *22*(6), 466–470. https://doi.org/10.1177/0963721413492764

Zaki, J., Ochsner, K. N., & Ochsner, K. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature Neuroscience*, *15*(5), 675–680. https://doi.org/10.1038/nn.3085

Zemla, J. C., & Austerweil, J. L. (2018). Estimating semantic networks of groups and individuals from fluency data. *Computational Brain & Behavior, 1*, 36–58.

Zemla, J. C., Cao, K., Mueller, K. D., & Austerweil, J. L. (2020). SNAFU: The semantic nework and fluency utility. *Behavior Research Methods*, *52*, 1681–1699.

Zerubavel, N., Bearman, P. S., Weber, J., & Ochsner, K. N. (2015). Neural mechanisms tracking popularity in real-world social networks. *Proceedings of the National Academy of Sciences*, *112*(49), 15072–15077. https://doi.org/10.1073/pnas.1511477112

Zhao, W., Feng, T., & Kazinka, R. (2014). The extensibility of the endowment effect to others is mediated by degree of intimacy. *Asian Journal of Social Psychology*, *17*, 296–301. https://doi.org/10.1111/ajsp.12061

Zhou, J., Gardner, M. P. H., Stalnaker, T. A., Ramus, S. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum, G. (2019). Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of value and task structure in an odor sequence task. *Current Biology*, *29*(6), 897-907.e3. https://doi.org/10.1016/j.cub.2019.01.048