MULTIDETERMINED CONCERN: ROLES FOR AFFECTIVE AND COGNITIVE PROCESSES IN EMPATHY AND ALTRUISM

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Psychology

By

Kristin Mary Brethel-Haurwitz, M.A.

Washington, DC
October 21, 2016
MULTIDETERMINED CONCERN: ROLES FOR AFFECTIVE AND COGNITIVE PROCESSES IN EMPATHY AND ALTRUISM

Kristin Mary Brethel-Haurwitz, M.A.

Thesis Advisor: Abigail A. Marsh, Ph.D.

ABSTRACT

Altruistic responding to distress relies on a series of empathic processes. These include perception of distress and vulnerability, affective resonance with this distress, and motivation to do something to improve the distressed individual’s well-being. Each of these components is independently critical for the production of effective prosocial helping, and failure at any step in this empathic process model will limit helping. Together, these processes are thought to result in a multidetermined concern, in which both bottom-up affective and top-down cognitive processes contribute to empathic and altruistic responding. How each of these processes contributes to particularly costly altruistic behavior remains to be determined. In three studies, the roles of affective and cognitive mechanisms at three steps in an empathic process model were tested with regard to their role in promoting costly altruism. Across all studies, empathic responding and prosocial behavior were examined in a population of extraordinary altruists – non-directed kidney donors. In Study 1, the role of the neural network supporting mammalian offspring care was examined in caring responding to cues of distress and vulnerability in altruists and matched controls. Results reveal that amygdala-mediated distress sensitivity and amygdala connectivity with the periaqueductal gray set altruists apart. In Study 2, neural resonance across the first-hand experience of pain and the observation of pain in a stranger, including both group-level overlap and individual-level correspondence in neural processing, was examined in altruists and controls. Altruists were found to have greater neural resonance supported by both group-level and
individual-level correspondence in left anterior insula. In Study 3, the role of empathic emotion regulation in promoting prosocial motivation and costly donation behavior was tested, first in a community sample and then in altruists and controls. While instructed reappraisal was found to affect donation behavior in Study 3a, especially for those high in trait reappraisal, effects of reappraisal in Study 3b were limited with regard to donation outcomes, suggesting that altruists and controls may instead differ in interacting effects of negative and positive affect on donation behavior across conditions. Implications for these interdependent processes in promoting altruism, and remaining questions, are discussed.
ACKNOWLEDGMENTS

This dissertation is the result of much support, encouragement, and mentorship over the last five years and beyond.

First and foremost, I am grateful to my advisor, Professor Abby Marsh. Through your mentorship, I have grown as a thoughtful, careful, and confident researcher and teacher. You have always taught me to aim high and take both successes and failures in stride, learning from each experience to become a better scientist. I am thankful for your time and support, and also the pushes to go beyond my comfort zone. At the end of my graduate career, I realize just how much I have learned from your mentorship. I feel so fortunate to have been a part of your lab and its incredible research. I look forward to continuing to work together for years to come.

This research would also not have been possible without the wonderfully supportive Laboratory on Social and Affective Neuroscience. I am so grateful to have had this team to work with to bring projects to life, solve problems big and small, face both rejections and successes, and commiserate about the realities of research. Each of you has made me a better researcher, and I am lucky to have you as colleagues. Thank you to Leah, Andrew, Joana, Emily, Kruti, Katie, Abbey, Lydia, and Elise, my partner in crime for the last five years. We are going to have to figure out how to do research without being a few feet away from each other at all times, but I am so excited to see what the future holds for each of us. I am also indebted to the many research assistants who have volunteered their time over the years, most importantly Maria, whose diligence helped this dissertation come to fruition. I am also grateful to all of the participants who have contributed their time and effort to our research, making all of this possible.

Beyond being a member of an incredible lab, I am fortunate to have been part of a fantastic Department. The faculty, staff, and students at Georgetown have created a wonderfully enriching environment for graduate students, and I am very grateful to have gone through this doctoral program. I am particularly thankful to my cohort, Anna, Erika, Nicole, and Laura, for their support in figuring out graduate school together. You are valued friends and colleagues.

I am grateful to my committee, Professor Yulia Chentsova Dutton, Professor John VanMeter, and Professor Jim Coan. I have sincerely appreciated your feedback and perspectives on research and teaching.

Most importantly, I am grateful for the support of my family. I am forever thankful to my parents, for always supporting my education and interests. I know that I am where I am now because asking questions and seeking answers was instilled as important and valuable early on. I have also had the encouragement of my sister, Meghan, even though she probably thinks this whole thing is pretty nerdy. I am particularly fortunate to have had my husband, Aaron, by my side since I first started studying psychology 11 years ago. Your unwavering support and encouragement have sustained me. You have kept life in perspective while always pushing me to be the best that I can be, in whatever it is that I want to pursue. Through Aaron I have also gained an additional support group in Linda, Ralph, Rachel, and Alex, who have always been eager to hear my research updates and encourage me along the way.
# TABLE OF CONTENTS

**INTRODUCTION** .............................................................................................................................. 1  
Empathy ............................................................................................................................................. 2  
Empathic Simulation ............................................................................................................................. 5  
Multidetermined Concern ..................................................................................................................... 8  
Goals of the Dissertation ...................................................................................................................... 9

**STUDY 1: AMYGDALA-MIDBRAIN CONNECTIVITY INDICATES A ROLE FOR THE MAMMALIAN PARENTAL CARE SYSTEM IN EXTRAORDINARY ALTRUISM** ........................................ 11  
Methods ............................................................................................................................................ 14  
Results ............................................................................................................................................... 23  
Discussion .......................................................................................................................................... 29

**STUDY 2: ALTRUISTS DEMONSTRATE ENHANCED SELF-OTHER OVERLAP IN PAIN PROCESSING IN ANTERIOR INSULA** ................................................................. 34  
Methods ............................................................................................................................................ 38  
Results ............................................................................................................................................... 47  
Discussion .......................................................................................................................................... 54

**STUDY 3: EMPATHIC EMOTION REGULATION** ............................................................................... 60  
Study 3a ............................................................................................................................................. 66  
Methods ............................................................................................................................................ 66  
Results ............................................................................................................................................... 74  
Study 3b ............................................................................................................................................. 81  
Methods ............................................................................................................................................ 81  
Results ............................................................................................................................................... 83  
Discussion .......................................................................................................................................... 91

**GENERAL DISCUSSION** ................................................................................................................ 96  
Conclusions ........................................................................................................................................ 101

**APPENDIX** ................................................................................................................................... 102

**REFERENCES** ................................................................................................................................. 104
LIST OF FIGURES

Figure 1.1. Task structure of implicit facial expression portion of task ........................................ 19
Figure 1.2. Left amygdala ROI ........................................................................................................ 24
Figure 1.3. Whole-brain results. ....................................................................................................... 25
Figure 1.4. Functional connectivity results ....................................................................................... 26
Figure 1.5. Structural connectivity in tracts linking amygdala and periaqueductal gray .............. 28
Figure 2.1. Trial structure of the empathic pain task ....................................................................... 42
Figure 2.2. Final subjective “slightly intense” pain level selected by each group ......................... 44
Figure 2.3. Other pain results .......................................................................................................... 48
Figure 2.4. Self pain results. ............................................................................................................ 49
Figure 2.5. Self-other correlation in BOLD activation in left anterior insula ROI ......................... 51
Figure 2.6. Perceptions of closeness as rated on the Inclusion of Other in the Self Scale. ............... 52
Figure 3.1. Task structure for the empathic emotion regulation task. ........................................... 70
Figure 3.2. Block structure for empathic emotion regulation task. ................................................ 71
Figure 3.3. Donation rates by condition in Study 3a. ..................................................................... 75
Figure 3.4. Correlations between trait reappraisal and donation rates by condition in Study 3a. .. 76
Figure 3.5. Correlations between positive and negative affect and donation rates by condition. ... 78
Figure 3.6. Affect intensity ratings across conditions in Study 3a. ................................................ 79
Figure 3.7. Index of affective arousal across conditions in Study 3a. ............................................. 80
Figure 3.8. Donation rates by group and condition in Study 3b. .................................................... 84
Figure 3.9. Correlations between affect and donation rates in each condition, by group ............. 87
Figure 3.10. Affect ratings across conditions for altruists and controls. ....................................... 89
Figure 3.11. Index of affective arousal across conditions for altruists and controls. ..................... 90
LIST OF TABLES

Table 1.1 Study 1 Participant Characteristics ................................................................. 15
Table 1.2 Example Distractor and Care-Eliciting Scenarios ............................................. 18
Table 1.3 Regions of Activation Predicted by Sympathy Difference Scores in
   Altruists > Controls ........................................................................................................ 24
Table 1.4 Functional Activation and Connectivity Results in Altruists > Controls,
   Without Subjects Exceeding Motion Parameters ......................................................... 27
Table 2.1 Study 2 Participant Characteristics .................................................................. 39
Table 2.2 Regions of Significant Activation for Other Pain > Other No Pain, Masked by
   Self Pain > Self No Pain in Each Group ...................................................................... 48
Table 2.3 Regions of Significant Activation for Self Pain > Self No Pain in Each Group .... 50
Table 2.4 Ratings of Intensity of Subjective Pain and Pain Unpleasantness by Group ....... 53
Table 2.5 Correlations between Objective Pain, Subjective Pain and Pain Unpleasantness,
   and ROI Mean Parameter Estimates for Self Pain > Self No Pain ................................ 54
Table 3.1 Affect Ratings for Task Photos from mTurk Validation .................................... 69
Table 3.2 Bivariate Correlations between Donation Rates and Self-Reported Affect
   in Study 3a ...................................................................................................................... 77
Table 3.3 Study 3b Participant Characteristics .................................................................. 82
Table 3.4 Bivariate Correlations between Donation Rates and Self-Reported Affect
   in Study 3b ...................................................................................................................... 86
Table 3.5 Study 3b Trait Measures and State Mood ......................................................... 88
Table A.1 Sympathy Ratings by Scenario Type in the Validation Task ............................ 103
INTRODUCTION

As a firefighter in Pierce County, Washington, it is Jo Kummerle’s job to save lives. What is different about Jo’s story is that she has saved lives while off duty as well. Several years ago, her mother was experiencing kidney failure and needed a life-saving kidney transplant. Jo was tested as a potential donor, but was not a match. Jo’s sister was able to participate in a transplant chain that helped their mother, but Jo still felt moved to help. So, she sought out an opportunity to donate one of her own kidneys to a total stranger through an altruistic donation. Through her altruism, Jo was able to save the lives of six others, since she kicked off a transplant chain herself. How did Jo make this jump from wanting to donate to her mother to donating to a complete stranger? How can we explain acts of such extraordinary altruism? This dissertation will focus on a set of processes hypothesized to be important for promoting empathy and altruism: sensitivity to distress and vulnerability, affective resonance, and empathic emotion regulation.

Altruism is characterized by behavior that is costly to the actor yet beneficial to the recipient (de Waal, 2008; Preston, 2013). There is evidence of altruism across social species, from rats (Ben-Ami Bartal, Decety, & Mason, 2011; Ben-Ami Bartal, Rodgers, Bernardz Sarria, Decety, & Mason, 2014) and bats (Wilkinson, 1984) to chimpanzees (Warneken, Hare, Melis, Hanus, & Tomasello, 2007), suggesting an adaptive value of this seemingly selfless category of behaviors. In humans, altruism is thought to stem from our complex cooperative tendencies (Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012), including shared offspring care (Hrdy, 2011). Importantly, altruism can have several different motives with distinct neurocognitive mechanisms (Marsh, 2016). Altruism based in reciprocity is dependent on expected return benefits for the actor, and is characterized by neural activation in structures associated with
reward processing (Hein, Morishima, Leiberg, Sul, & Fehr, 2016; Rilling & Sanfey, 2011; Rilling et al., 2002; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). Kin-based altruism can also be reciprocal in nature, but is additionally driven by the indirect fitness benefits of supporting the survival of genetically related others. Such kin selection relies on limbic and cortical structures that support the recognition of familiar others (Lieberman, Tooby, & Cosmides, 2007; Platek & Kemp, 2009; Platek, Keenan, & Mohamed, 2005). In contrast, care-based altruism is dependent on empathic processes including distress perception, affective resonance, and concerned approach motivation (Marsh, 2016; Preston, 2013). The donation of a kidney to a stranger meets even the most stringent definitions of human altruism, in that it is a behavior that is costly to the donor while benefitting an anonymous, non-kin other (Batson, 2010; Clavien & Chapuisat, 2013; de Waal, 2008), thus eliminating motives based in reciprocity or kin selection. As such, altruistic kidney donors were a focus in examining empathic processes that drive care-based altruism in the current dissertation.

**Empathy**

Empathy relies on interpersonal resonance, consistent with its translation to English from the German *einfühlung*, meaning “feeling into” (Titchener, 1909). Whether we intend to or not, through empathy we can quickly internalize others’ emotional experiences, simply by observing their circumstances or behavior. It is thought that humans are particularly empathically adept due to our evolutionary history that required cooperation among individuals to survive, and specifically the shared care of offspring known as alloparenting (de Waal, 2008; Decety, Norman, Berntson, & Cacioppo, 2012; Preston, 2013). Such interdependent existences (Tomasello et al., 2012) meant that our ancestors had to quickly pick up on cues communicating the needs and intentions of others, feel motivated to act, and respond appropriately.
Empathy is composed of the suite of processes through which we connect with others. Empathic processes range from rapid, low-level physiological and neural resonance to effortful and abstract cognitive processes to understand another’s context and perspective. Ways of defining empathy and its subcomponents are numerous, leading to conceptual inconsistency across the empathy literature. Many authors have divided empathy into affective and cognitive components (Bird & Viding, 2014; Blair, 2005; Dvash & Shamay-Tsoory, 2014; Shamay-Tsoory, 2011). Further dividing affective empathy, de Waal (2008) and others (e.g., Decety, 2015) categorize empathic processes into a hierarchy of three levels, ranging in cognitive demand and complexity: emotion contagion, empathic concern, and perspective taking. In contrast to the more affective components of emotion contagion and empathic concern, the more cognitively complex process of perspective taking, in which people understand others’ contexts and imagine themselves in their shoes, is not sufficient for prosocial responding (Blair, 2008; Jones, Happe, Gilbert, Burnett, & Viding, 2010; Nichols, 2001). As such, affective processes have been implicated as more central to empathy associated with prosocial behavior and altruism, and thus will be a focus of this dissertation.

Within this three-tiered framework, emotion contagion is empathy’s most basic building block. As its name suggests, emotion contagion is the low-level process by which we catch another’s emotional state (de Waal, 2008). It can occur outside conscious awareness and without top-down input, though top-down modulation of emotion contagion is possible. Importantly, sharing in another’s emotional experience does not presuppose an understanding, or even an awareness, of the contextual details surrounding the emotional experience of the other individual, nor does it assure concern for another’s well-being, which begins at the next level of the empathic hierarchy (de Waal, 2008; Decety, 2015). In empathic concern, one may not only share
in another’s emotional experience, but also gain a conscious awareness of that person’s state and concern for his or her welfare, leading to a desire to improve his or her well-being. Empathic concern has also commonly been called sympathy or sympathetic concern (e.g., Eisenberg et al., 1989a; Eisenberg et al., 1989b). In this way, empathy and sympathy are not necessarily synonymous – while the larger construct of empathy can also include emotion contagion and affective or cognitive perspective taking, which may occur independently of empathic concern, sympathy is specific to the concern felt for others and their welfare. Empathic concern is closely related to compassion, which is thought to be a specific emotion orientation toward the distress of another, distinct from personal distress, associated with a caring attitude and prosocial approach toward the distressed target (Goetz, Keltner, & Simon-Thomas, 2010). As such, empathic concern has been positively linked with prosocial helping behaviors and negatively linked with antisocial harming behaviors (Eisenberg & Fabes, 1990; Miller & Eisenberg, 1988; Vachon, Lynam, & Johnson, 2014).

Nichols (2001) proposes a “concern mechanism” driven by our emotional resonance with a distressed target as providing the motivation to engage in helping behaviors, which does not rely on extensive cognitive perspective taking. Instead, we only need to be able to differentiate between self and other (Nichols, 2001). Nichols also emphasizes distress recognition as an important prerequisite for empathic concern, in which one must first be consciously aware of distress in another, and accurately recognize it as such, before being moved by it. This argument is supported by findings that accurate recognition of fearful facial expressions predicts prosocial behavior in the laboratory (Marsh, Kozak, & Ambady, 2007). This also requires that a target be sufficiently expressive for an observer to pick up on their distress and thus develop empathic concern (Zaki, Bolger, & Ochsner, 2008).
Linking together these empathic processes, Decety and colleagues (Decety, 2010; Decety, Bartal, Uzefovsky, & Knafo-Noam, 2016) suggest that the mechanisms by which empathy leads to prosocial helping behavior can be conceptualized as a process model. This model links together the affective components of empathy, emotion contagion and empathic concern, and also other emotional processes that support affective empathy in producing concerned responding, including distress perception and emotion regulation. In this model, first distress must be perceived. Then, depending on the social context, this may lead to affective resonance with the distressed individual, in which the observer experiences aversive negative affect and associated neural activation. Depending on the extremity of affective resonance and the ability to regulate such empathic reactivity, prosocial approach or self-interested avoidance motivation will result. If the observer perceives an effective ability to help without prohibitive cost, helping behavior will follow, which may then be reinforced as socially rewarding. This process model is not entirely novel, but rather brings together processes that have separately been tied to prosocial outcomes, and focuses on those that are expected to predict helping across species, even including helping behavior in rats (Ben-Ami Bartal et al., 2011; Ben-Ami Bartal et al., 2014), suggesting evolutionarily conserved mechanisms including those adapted for offspring care (Decety et al., 2016).

**Empathic Simulation**

According to this model, accurate recognition of emotions in others and affective resonance are important prerequisites for prosocial behavior. In order for prosocial behavior to be initiated, an observer must perceive that a target is in distress and also be motivated to see the target’s welfare improved. A failure at either of these steps would limit the possibility of prosociality. Several major theories of the mechanisms supporting empathic processing suggest
that we achieve empathic resonance with others via physiological simulation of their states (de Vignemont & Singer, 2006; Decety & Jackson, 2004). Such theories are based in a more general theory of mirrored processing of perceptions and actions (Prinz, 1997), but have been extended to corresponding emotional states. Preston and de Waal (2002) propose an empathic Perception Action Model in which the sensory and emotional experiences of others are directly mirrored in our own physiology. This model suggests that when we observe others anticipating or experiencing pain, we not only generate a representation of their fear and discomfort, but the corresponding autonomic and neural signatures of fear and pain are activated. Support for simulation models of empathy comes from studies of neural and peripheral physiology in laboratory participants, particularly with regard to experienced and observed pain and fear. Specifically, similar activation for first-hand and vicarious distress in pain-related regions including the anterior insula and mid-cingulate cortex and fear-relevant structures like the amygdala has been found across studies (Britton, Taylor, Sudheimer, & Liberzon, 2006; Lamm, Decety, & Singer, 2011).

Costly altruism such as altruistic kidney donation presents a potential paradox – such altruists are affectively sensitive to the distress of others, demonstrating heightened amygdala activation in response to fearful facial expressions (Marsh et al., 2014), yet engage in a potentially risky major surgery to help a complete stranger. Many would find such an act, even when directed toward a close other, to be highly distressing in itself. How might altruistic kidney donors balance their concern for the recipient with their concern for their own well-being?

While feeling into the emotions of another makes their distress salient and meaningful, and contributes to the motivation to alleviate their distress, it also has the potential to create an aversive state that leads to avoidance. With excessive personal distress, attention and motivation
turn inward, toward decreasing negative affect and deficits in well-being, rather than directly addressing distress in another. Thus, instead of engaging in prosocial approach toward a distressed target, someone experiencing significant personal distress via affective resonance may avoid the source of that distress, which in this case is the needy other. Such a balance between personal distress and empathic concern has been the focus of an ongoing debate regarding the motivations for prosocial behavior.

Recent research examining the neural correlates of empathic processes also supports this conclusion, in that empathic concern is more likely than personal distress to predict costly prosociality, which is associated with neural activation in mesolimbic areas thought to be associated with approach motivation (FeldmanHall, Dalgleish, Evans, & Mobbs, 2014). Batson and colleagues have studied this balance between personal distress and empathic concern in a series of studies in which escape from a distressed target is offered as an alternative to helping. Given the balance between costs and benefits of prosocial behavior, it is expected that if a less costly means for decreasing one’s own personal distress, such as avoidance of the distressed target, is possible, then prosocial helping will not occur. In several studies, they found that those higher in personal distress were less likely to help when escape from the need for help was easy, but this was not the case for those higher in empathic concern (Batson, Fultz, & Schoenrade, 1987). Further, if feelings of empathic concern were induced, the tendency to seek easy escape was greatly reduced (Batson et al., 1987). Such a balance between empathic concern and personal distress, in which concern is more likely than distress to lead to prosocial outcomes, has been discussed extensively by Eisenberg and colleagues, particularly in children (Eisenberg et al., 1994; Eisenberg, 2000).
Multidetermined Concern

Vaish (2016) argues that flexible empathic responding depends on a combination of both bottom-up and top-down processes, resulting in a multidetermined concern. While affective resonance is expected to be largely reflexive and bottom-up, more reflective top-down processes including self-other differentiation and emotion regulation then modulate this resonance. Concern is multidetermined, in that one responds to distress whether it is in the form of a facial expression, sound, semantic description, abstract cue, and so on. Empathic responding is flexible in that it can be modified by the context, in which one may have greater empathic resonance with close (Singer et al., 2004), similar (Hein, Silani, Preuschoff, Batson, & Singer, 2010), or preferred (e.g., fair) (Singer et al., 2006) others. One such top-down process that could modulate a more bottom-up affective arousal is emotion regulation. Emotion regulation encompasses a set of strategies through which people can modulate their emotional reaction to a situation (Gross, 1998b). The application of emotion regulation strategies to empathic emotions parallels the application of emotion regulation to intrapersonal emotion. With regard to regulating empathic distress, emotion regulation is still a goal-directed, motivated state, in which the goal is to modify the emotional state (which now originates from the experience of another), which could have consequences for prosociality, leading to either help or avoidance of the needy target. And, this motivation stems from a discrepancy between a current state and a desired state, in which people may desire to be more or less emotionally affected by another’s distress, and their own level of distress may result from the degree of discrepancy between the needy target’s current state and goals. Further, as with intrapersonal emotion regulation (Mauss, Bunge, & Gross, 2007), empathic emotion regulation can be either explicit or implicit, occurring as a result of
conscious and deliberate efforts, or occurring outside of awareness and without a conscious goal to regulate.

The regulatory processes that we apply to our intrapersonal emotions can also be applied to the interpersonal emotions we experience as a result of the emotional circumstances of others (Zaki, 2014). Decety (2010) has argued that along with emotion understanding and empathic arousal, emotion regulation is a core component of empathic processing that may successfully translate to prosocial helping, while Eisenberg and colleagues (Eisenberg & Okun, 1996; Eisenberg et al., 1994; Eisenberg et al., 1996; Eisenberg, 2000; Eisenberg, Wentzel, & Harris, 1998; Fabes, Eisenberg, Karbon, Troyer, & Switzer, 1994) have consistently suggested the development of self-regulation is critical for the development of empathy and prosocial tendencies. In particular, they hypothesized that while personal distress and sympathy (their term for empathic concern) would both be predicted by emotional reactivity, emotion regulation would positively predict sympathy while negatively predicting personal distress (Eisenberg et al., 1994). They confirmed these hypotheses empirically in adults, finding that dispositional high reactivity and low regulation predicted personal distress, while moderate reactivity and regulation (when controlling for reactivity) predicted sympathy (Eisenberg et al., 1994).

**Goals of the Dissertation**

The goals of this dissertation are to answer three major questions regarding affective and cognitive processes that support empathy and altruism. First, does the amygdala-mediated distress sensitivity previously observed in extraordinary altruists contribute to caring decision-making? In Study 1, a potential role for the mammalian offspring care system was tested by examining neural responses to cues of distress and vulnerability in altruists versus matched controls. Second, in addition to enhanced sensitivity to distress in others, do altruists also
demonstrate enhanced similarity in first-hand and vicarious experiences of pain? Study 2 addressed this question with an empathic pain paradigm in which neural responses to first-hand and vicarious pain were compared in altruists and matched controls to examine if altruists might be characterized by a heightened neural empathic resonance. Third, are emotion regulatory processes important for promoting well-regulated empathic concern and prosocial helping? Studies 3a and 3b were designed to test the effects of two potentially contradictory cognitive reappraisal strategies in regulating empathic affect. In these studies, an approach-oriented hopeful appraisal was contrasted with an avoidance-oriented distancing appraisal in influencing costly helping behavior toward strangers. Together, these studies suggest that both bottom-up affective processes and top-down cognitive processes have roles in empathic responding and consequent prosocial and altruistic behavior. Each study represents a step forward in understanding the mechanisms that support the extension of concern and care beyond kin in humans. The implications of these findings, and also the questions raised for future research, are discussed in the General Discussion.
STUDY 1: AMYGDALA-MIDBRAIN CONNECTIVITY INDICATES A ROLE FOR THE MAMMALIAN PARENTAL CARE SYSTEM IN EXTRAORDINARY ALTRUISM


Why would an individual ever intentionally risk his or her life to benefit a stranger? Risky, life-saving acts on behalf of strangers represent extraordinary manifestations of altruism, broadly defined as a voluntary, costly behavior aimed at improving the welfare of another individual (Batson, 2010; de Waal, 2008). Altruism can assume several distinct forms, each reflecting distinguishable underlying neurocognitive processes (Burkart et al., 2014; de Waal, 2008; Hein et al., 2016; Marsh, 2016). Costly altruism toward strangers, while rare, has been documented across multiple social species, including humans, chimpanzees, and dolphins (de Waal, 2008; de Waal, 2012; Preston, 2013) and is thought to reflect the exploitation of deep and evolutionarily conserved proximal neural mechanisms that evolved to support parental care for distressed or vulnerable young (Marsh, 2016; Preston, 2013). At minimum, these mechanisms must support two capacities: to encode cues that signal distress and vulnerability, and to bias the organism toward protective rather than preservative behaviors in response (Preston, 2013).

Recent research has implicated the structure and function of the amygdala in costly altruism (Marsh et al., 2014). The amygdala is a subcortical neural structure that can rapidly encode nonverbal distress cues (Fusar-Poli et al., 2009; Sprengelmeyer et al., 1999) and is thought to be the point of entry to the parental care system (Preston, 2013). Coordinated transmission of information from the amygdala to subcortical regions including the striatum, bed nucleus of the stria terminals, various nuclei of the hypothalamus, and periaqueductal gray support the generation of caregiving motivation and behavior following the detection of distress.
in juveniles or otherwise vulnerable individuals (Nephew, Caffrey, Felix-Ortiz, Ferris, & Febo, 2009; Numan, 2006; Preston, 2013). Consistent with this, amygdala-mediated fear sensitivity corresponds both to individual variation in altruistic and empathic responding in humans (Desbordes et al., 2012; Lutz, Brefczynski-Lewis, Johnstone, & Davidson, 2008; Marsh et al., 2014) and to individual variation in parental and alloparental care in other mammalian species (Bosch, 2013; Martel, Nishi, & Shumyatsky, 2008; Toscano, Bauman, Mason, & Amaral, 2009); among primate species, provision of alloparental care is the single strongest predictor of the frequency of altruistic helping of adults (Burkart et al., 2014).

When distress is detected, generating an altruistic response may be particularly dependent on reciprocal connections between the amygdala and periaqueductal gray (Almeida, Giovenardi, Silva, Oliveira, & Stein, 2006), an integrative midbrain region that curls around the cerebral aqueduct and supports various basic survival-related functions that include defensive behaviors, pain modulation, anxiety, and reproductive behaviors such as parental care (Behbehani, 1995; Linnman, Moulton, Barmettler, Becerra, & Borsook, 2012). Forebrain projections to the periaqueductal gray arise primarily from the central nucleus of the amygdala—which also receives reciprocal projections from this region—as well as from the hypothalamus, insular cortex, and prefrontal cortex (Linnman et al., 2012; Rizvi, Ennis, Behbehani, & Shipley, 1991). Periaqueductal gray plays an essential role in regulating behavioral inhibition in response to amygdala-signaled threat cues (Preston, 2013; Tovote et al., 2016). This is critical to altruistic responding, which requires overriding defensive responses to amygdala-mediated detection of distress. Rodent studies find that ablations of various sites within periaqueductal gray impair facets of normal caregiving behavior (Lonstein & Stern, 1997; Lonstein & Stern, 1998; Lonstein, Simmons, & Stern, 1998). Observationally, activity in periaqueductal gray (and amygdala)
increases along with protective behavior in female mice whose pups are under threat (Nephew et al., 2009), and in human mothers exposed to infant cries or images (Swain, 2011). In humans, activity in this region also increases following induction of compassion or caregiving motivation toward vulnerable or distressed unknown individuals (Decety, Echols, & Correll, 2010; Kim et al., 2009; Simon-Thomas et al., 2012). Conversely, unusually low levels of compassion are linked to reduced periaqueductal gray responsiveness to perceived distress (Decety, Skelly, & Kiehl, 2013; Michalska, Zeffiro, & Decety, 2016).

We therefore hypothesized that extraordinary acts of altruism may result from enhanced care-oriented responses to distress that are supported by recruitment of amygdala-midbrain circuitry. To test this hypothesis, we evaluated activity and connectivity in this circuit in a sample of extraordinary altruists and matched controls as they responded to depictions of others’ distress (Marsh & Ambady, 2007). Altruists had all donated a kidney to a stranger with whom they had no prior relationship. Costs of these donations include extensive preoperative screening, significant postoperative pain, small risks to donors’ survival and long-term health, and financial sacrifices (Rodrique et al., 2016). In part due to these challenges, such donations are extremely rare, with fewer than 2,000 such donations recorded in the United States through the end of 2015 (data from Organ Procurement and Transplantation Network). Whereas previous findings have identified enhanced amygdala reactivity in this population during passive viewing of fearful facial expressions (Marsh et al., 2014), the present research aimed to directly link caring decision-making in this population to caregiving-associated neural circuits through the use of functional magnetic resonance imaging (fMRI) and structural MRI, including diffusion tensor imaging (DTI). Because altruistic responses to distress are thought to be rapid and intuitive (Rand & Epstein, 2014), we predicted that enhanced altruistic responding would correspond to
elevated responsiveness and functional coupling in amygdala and periaqueductal gray following even preattentively presented nonverbal distress cues, and this pattern would be supported by enhanced structural connectivity between these regions.

**Methods**

**Participants**

Thirty-nine healthy adults between 23 and 56 years old (Table 1.1) took part in this study for monetary payment. Nineteen altruistic kidney donors (7 women) were recruited using mailings and electronic advertisements through local and national transplant organizations. The sample of altruists was limited by the extreme rarity of this behavior. Only 1,265 adults had ever directed a kidney to a stranger in the United States through the end of 2012, when recruitment was completed (0.0005% of the adult U.S. population in 2012; data from the Organ Procurement and Transplantation Network). Because altruists were recruited from across North America, most altruists resided more than a two-hour drive from the university and were provided with airfare and up to two nights’ lodging. All altruists had donated a kidney to a stranger unknown to them personally at the time of donation. Sixteen altruists were non-directed donors for whom the recipient was anonymous at the time of donation. The remaining three directed their donations to a specific individual who was known to them at the time of donation but whose need for a kidney they had learned about through, for example, a flier or Internet posting. All donations were verified through independent sources, including transplant center records or media reports. Using data obtained from the Organ Procurement and Transplantation Network, we confirmed that altruists recruited for this study were representative of the national population of altruistic donors at that time in terms of sex and race (exact ages are not available for the national sample).
addition, 20 healthy volunteers (11 women) were recruited from the local community using fliers, online advertisements, and electronic participant databases including ResearchMatch.

Table 1.1

Study 1 Participant Characteristics

|                          | Altruists (n = 19) | Controls (n = 20) | p  
|--------------------------|--------------------|------------------|------
| Male/Female (% Male)     | 12/7 (63.16%)      | 9/11 (45.00%)    | .256
| White/Other Race (% White) | 18/1 (94.74%)   | 17/3 (85.00%)    | .316
| Right/Left Handed (% Right) | 18/1 (94.74%) | 19/1 (95.00%)    | .998
| Household Income ≥ $60,000 | 13 (68.42%)       | 8 (40.00%)       | .268
| Education ≥ Four-Year Degree | 12 (63.16%)  | 16 (80.00%)      | .243
| Age M (SD)               | 46.32 (8.68)       | 44.75 (6.41)     | .524
| IQ M (SD)                | 115.74 (11.14)     | 111.95 (13.06)   | .338

Note. Four controls did not report their household income.

Exclusion criteria for all participants included current use of psychotropic medication, history of head injury or neurological illness, IQ < 80 (as assessed using the Kaufman Brief Intelligence Test - Second Edition; (Kaufman & Kaufman, 2004), and pregnancy or other contraindications to safe MRI scanning, including metal fragments or implants. Controls were excluded if they reported having ever volunteered to donate an organ to any individual (not including consenting to become a deceased organ donor). All study procedures were approved by the Internal Review Board at Georgetown University in Washington, DC, and all participants provided written informed consent before testing.

Two altruists and one control were excluded from functional MRI analyses due to a computer monitor error that affected the timing of the preattentive facial expressions. This resulted in an fMRI subsample of 35 participants, 17 of whom were altruists and 18 of whom were matched controls. Within this subsample, altruists and controls did not differ on demographic variables reported for the full sample in Table 1.1, all p > .05. These three excluded
participants were retained in the DTI analyses, as was one control who did not complete the fMRI task, yielding the full sample of 39 participants.

Procedures

All interested volunteers initially completed a 90-minute online survey assessing exclusion and inclusion criteria and demographic variables. Online screening also assessed self-reported empathy via the Interpersonal Reactivity Index (IRI; Davis, 1983). This questionnaire measures four components of empathy via four subscales that assess both affective empathy (Empathic Concern, Personal Distress) and cognitive empathy (Perspective Taking, Fantasy). Eligible volunteers were then screened by telephone to confirm eligibility. Researchers coordinated altruists’ travel to and lodging at Georgetown University to enable on-site neuroimaging and cognitive testing. To ensure groups were matched, eligible controls completed laboratory screening that included assessments of IQ, income, education, psychological history, medication use, and MRI compatibility before MRI scanning. After confirmation of eligibility, controls completed neurocognitive tasks and MRI scanning in a final visit.

Neuroimaging Acquisition and Task

MR images were acquired with a 3T Siemens Tim Trio scanner (Siemens Medical Solutions) and a 12-channel phased-array head coil. Functional data were collected using a T2*-weighted echo-planar imaging sequence (46 3.0 mm transversal slices; 64 × 64 matrix; repetition time, 2,500 ms; echo time, 30 ms; field of view, 192 mm²; 3.0 × 3.0 × 3.0 mm voxels). The first four volumes of each functional run were excluded from analysis to account for magnet stabilization. High-resolution T1-weighted anatomical images were also acquired (3D Magnetization Prepared Rapid Acquisition Gradient Echo; 176 1.0-mm axial slices; field of view, 250 mm²; repetition time, 1,900 ms; echo time, 2.52 ms; 246 × 256 matrix). DTI data were
collected using two runs of an echo-planar pulse sequence (repetition time, 6300 ms; echo time, 86 ms; 2.5 x 2.5 x 2.5 mm voxels; diffusion directions, 30; b value, 1000 s/mm$^2$; number of b value = 0s/mm$^2$ images, 5).

Participants completed four randomized runs of the scenarios task, each lasting 8 minutes and 22 seconds. Each run featured presentations of 20 brief scenarios that each appeared for 9 seconds, and were either care-eliciting (i.e., featuring a protagonist who was the target of another individual’s aggressive or callous behavior) or not, with 10 of each scenario type included per run (Table 1.2). Scenarios were adapted from a previous behavioral version of this paradigm (Marsh & Ambady, 2007). Each scenario was followed by 3.5-second prompt in response to which participants reported their sympathy using a 1 (no sympathy at all) to 4 (a lot of sympathy) scale via button press. (This term was selected as more specific than “caring” and because self-reported sympathy during this task correlates at $r = .90$ with self-reported desire to help the protagonist (Marsh & Ambady, 2007)).

Each scenario was preceded by 4 facial expressions presented in the context of a lexical decision task (deciding whether a string of letters formed a real or nonsense word via button press), which participants were told was the focal task, following Marsh and Ambady (2007) (Figure 1.1). Word type was presented randomly. Affectively neutral letter strings were selected from the ANEW database (Bradley & Lang, 1999). Each string of letters was presented for 1.5 seconds, and was preceded by either a 27 ms fearful or neutral facial expression (Dimberg & Öhman, 1996), which was then immediately backward masked by an 80 ms scrambled neutral face. Fearful expressions were chosen as cues that signal both vulnerability and distress (Marsh, 2016). Facial expressions were 2 males and 2 females drawn from the Pictures of Facial Affect stimulus set (Ekman & Friesen, 1976). Participants completed 4 such expression-mask-letter
string sequences before each scenario. Within each block, all expressions were either fearful or neutral.

This task structure resulted in a total of 80 scenarios, 40 of which featured care-eliciting protagonists and 40 of which were distractor scenarios. Of each type of scenario, 20 were preceded by fearful expressions and 20 by neutral expressions. The full set of scenarios was validated in a separate sample of 55 participants, who reliably reported more sympathy in response to the care-eliciting scenarios than the distractor scenarios, $F(1, 54) = 256.39, p < .001$, and, consistent with prior results, reported the most sympathy when the care-eliciting scenarios were preceded by preattentively presented fearful expressions relative to neutral expressions, $t(54) = 11.46, p < .001$, corroborating the efficacy of combined depictions of vulnerability and distress in eliciting increased care (see Appendix). The combination of these cues in particular, rather than presenting distress and vulnerability in isolation, is thought to amplify caring responses (Preston, 2013). Presentation of the four scenario types was randomly ordered within each run. Jittered fixations, randomly varying from 1 to 4 seconds preceded and followed each scenario-question sequence. Each run concluded with a final 12-second fixation. Runs one and two were separated from runs three and four by a 10-minute DTI scan.

Table 1

Example Distractor and Care-Eliciting Scenarios

<table>
<thead>
<tr>
<th>Distractor Scenario</th>
<th>Care-Eliciting Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peter promised his mom that he would practice his guitar if she bought him one. She bought the guitar for Peter, but Peter never practiced because he wanted to play with his friends instead.</td>
<td>Karen is sitting at the school cafeteria eating her lunch. A group of girls start making fun of her weight. Karen just wanted to eat her lunch in peace, but the other girls kept making fun of her.</td>
</tr>
</tbody>
</table>
Analysis of Behavioral Data

Mean ratings of sympathy were calculated for the care-eliciting scenarios preceded by fearful versus neutral expressions. Then the difference between these scores (fearful > neutral expressions) was calculated to index the degree to which the fearful expressions elicited increased sympathy in individual participants. No group difference in these difference scores emerged, \( t(33) = 1.33, p = .193 \), (altruist \( M = 0.01, SD = 0.21 \); control \( M = -0.13, SD = 0.39 \)).

Self-reported empathy measured via the IRI was calculated for each of the four 7-item subscales. Total scores were calculated by summing the four subscales. Group means were compared via independent sample \( t \) tests. Groups did not differ in total or subscale self-reported empathy, all \( p > .05 \). Associations between sympathy ratings, self-reported empathy, and neural measures were tested via Pearson’s or Spearman rank correlations, as noted in the Methods and Results.
Analysis of Neuroimaging Data

**Average functional activation analyses.** Functional data were preprocessed and analyzed according to the general linear model, using Analysis of Functional NeuroImages (AFNI; Cox, 1996). The four runs of the task were concatenated, despiked, motion-corrected, and spatially smoothed using a 6 mm full-width half maximum Gaussian filter. Functional data were aligned to the anatomical grid, transformed to the Talairach and Tournoux Atlas (Talairach & Tournoux, 1988), and masked with an extents mask to account for motion artifacts and to exclude voxels without valid data at every TR for every run, helping to control for false activations. Ten regressors were created to model task events: fearful expression blocks, neutral expression blocks, care-eliciting scenarios preceded by fearful expressions, care-eliciting scenarios preceded by neutral expressions, distractor scenarios preceded by fearful expressions, distractor scenarios preceded by neutral expressions, and four regressors for question blocks (corresponding to each of the four preceding scenario types). Fixation trials were modeled implicitly; baseline was modeled by a first-order function, and motion artifacts were modeled using the six estimated rigid-body motion parameters. Boxcar regressors representing the occurrence of each block type were convolved with a canonical hemodynamic response function, scaled to an amplitude of 1.

Group-level analyses were limited to a brain mask defined by voxels with functional activation shared by at least 50% of participants. Cluster size thresholds were calculated for a corrected clusterwise p threshold of .05 for this group mask (whole brain clusterwise corrected p = .05: 350 contiguous voxels at uncorrected p = .05), and also anatomical left and right amygdala masks (left amygdala clusterwise corrected p = .05: 10 contiguous voxels at uncorrected p = .05; right amygdala clusterwise corrected p = .05: 9 contiguous voxels at uncorrected p = .05), using
10,000 Monte Carlo simulations conducted via 3dClustSim in AFNI (June 2016 AFNI 16.1.26 version of 3dClustSim, utilizing a Gaussian plus mono-exponential spatial autocorrelation function to decrease risk of false positives). Anatomical masks were defined using the Talairach-Tournoux Atlas in AFNI. MNI coordinates of peak t statistics within significant clusters are reported. Average group neural activation in whole-brain and region of interest (ROI) analyses were compared via independent sample t tests, with reported sympathy included as a covariate of interest in predicting neural activation, as described in Results.

**Functional connectivity analyses.** Generalized psychophysiological interaction (gPPI) analyses were conducted in SPM8 (Wellcome Trust Department of Cognitive Neurology) using the generalized PPI toolbox (McLaren, Ries, Xu, & Johnson, 2012). Functional images were slice-time corrected, realigned, coregistered to anatomical scans, normalized to $2.0 \times 2.0 \times 2.0$ mm voxel size in MNI space using parameters calculated during segmentation of anatomical scans, and smoothed using a 6 mm Gaussian kernel. Task-specific functional connectivity with the left amygdala was estimated using gPPI analysis (McLaren et al., 2012). The amygdala seed region was anatomically defined using the AAL atlas (Tzourio-Mazoyer et al., 2002). A design matrix was created for each participant that included the stimulus time series for each of the ten conditions and six motion parameters, which were convolved with a hemodynamic response function to create psychological regressors. A cluster size threshold was calculated for a corrected clusterwise $p$ threshold of .05 for a periaqueductal gray ROI defined as a 10 mm diameter sphere centered at the average coordinates for left and right periaqueductal gray activation ($x = \pm 4$, $y = -29$, $z = -12$) from a recent meta-analysis (Linnman et al., 2012) (left periaqueductal gray clusterwise corrected $p = .05$: 21 contiguous voxels at uncorrected $p = .05$; right periaqueductal gray clusterwise corrected $p = .05$: 21 contiguous voxels at uncorrected $p = \ldots$
.05), using 10,000 Monte Carlo simulations conducted via 3dClustSim in AFNI (June 2016 AFNI 16.1.26 version of 3dClustSim utilizing a Gaussian plus mono-exponential spatial autocorrelation function to decrease risk of false positives). These thresholds were applied to the ROI independent sample t tests conducted on the gPPI results.

**Structural connectivity analyses.** Probabilistic tractography between the amygdala and periaqueductal gray was used to investigate white matter integrity within subject-specific tracts. Standard preprocessing was performed using FMRIB’s Diffusion Toolbox (FDT) within the FSL software package (http://www.fmrib.ox.ac.uk/fsl/). Diffusion images were skull-stripped and underwent eddy current correction before being visually inspected for motion and image artifacts. Diffusion images were registered to T1 structural space using a linear transformation. Freesurfer volumetric segmentation was performed to obtain subject specific amygdala seeds, which were then transformed into diffusion space.

The periaqueductal gray target mask was drawn by hand on the B0 image for all subjects using Duvernoy’s atlas (Naidlich et al., 2009) and figures from Ezra and colleagues (Ezra, Faull, Jbabdi, & Pattinson, 2015) for anatomical reference. The rater was blind to subject group membership throughout drawing and tractography analysis. Probabilistic tractography with multiple fiber orientations (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007) was run separately for left and right amygdala seeds. To exclude erroneous voxels, exclusion masks were created to eliminate paths traveling: (1) rostral to the amygdala; (2) caudal to the periaqueductal gray; (3) dorsal to the periaqueductal gray; (4) through the hypothalamus (3D box with MNI coordinates: x = -10:10, y = -16:6, z = 0:20); and (4) through the stria terminalis (box with MNI coordinates: x = -22:22, y = -22:14, z = 4:22). Boxes were drawn in standard MNI space and transformed into diffusion space for analysis. Ten thousand samples per seed voxel were
modeled and results were normalized for the number of samples reaching the target periaqueductal gray mask (divided by waytotal) and thresholded at 10% to create subject-specific tracts. Subjects with fewer than 10 voxels meeting this criterion were excluded per hemisphere (one control was excluded from the left hemisphere analysis, one control was excluded from the right hemisphere analysis). Tracts were visually inspected for accuracy. Finally, dtifit was used to obtain an FA image, which was thresholded at 0.1 and masked with the subject-specific tract to obtain a mean FA within the amygdala-periaqueductal gray tract. All analyses took place in subject diffusion space, but were transformed to MNI space for visualizations. Given a non-normal distribution, non-parametric statistics were used to examine FA values. Spearman’s rank correlations were conducted to assess the relationship between FA values in the left amygdala-periaqueductal gray tract and total and subscale self-reported empathy on the IRI, after controlling for participant age and IQ.

Results

Functional activation and connectivity

We aimed to identify patterns of neural activation that are associated with enhancements in altruists’ caregiving motivation in response to distress and vulnerability. We conducted contrast tests within left and right amygdala ROIs to compare blood-oxygenation level dependent (BOLD) responses as participants responded to care-eliciting scenarios that were preceded by masked fearful versus neutral expressions (fearful > neutral expressions). We included participants’ reported sympathy ratings in response to the scenarios (fearful > neutral expressions) as a covariate to identify regions in which increased activity during this epoch corresponded specifically to increases in caring responses. These analyses were applied to anatomical masks of left and right amygdala, separately. Increasing sympathy for the protagonist
following preattentive fearful expressions corresponded to increased activation in the left amygdala in altruists relative to controls ($x = -27, y = -3, z = -22, t(32) = 3.73, p < .05_{corrected}$; Figure 1.2). No similar clusters emerged in the right amygdala. A corresponding whole-brain analysis was also conducted, which revealed that increasing sympathy predicted increased activation for altruists in a cluster of activation that extended from the midcingulate and posterior cingulate into periaqueductal gray, and clusters in bilateral dorsolateral prefrontal cortex, precuneus, bilateral amygdala, and bilateral insula, all $p < .05_{corrected}$ (Table 1.3, Figure 1.3). No significant clusters were preferentially associated with increased sympathy in controls.

Figure 1.2. Left amygdala ROI. 
(A) Sympathy difference scores predict greater neural activation in altruists than controls in left amygdala for care-eliciting scenarios preceded by fearful expressions versus neutral expressions. $y = -3. k = 22$. Color bar = $t$ value. Clusterwise corrected threshold $p < .05$. (B) Plotting the relationship between left amygdala activation during this epoch and sympathy difference scores confirmed a positive linear slope for altruists.

Table 1.3

<table>
<thead>
<tr>
<th></th>
<th>Peak x</th>
<th>Peak y</th>
<th>Peak z</th>
<th>Peak $t(32)$</th>
<th>Peak Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>4418</td>
<td>2</td>
<td>-32</td>
<td>62</td>
<td>5.96</td>
<td>Right Paracentral Lobule</td>
</tr>
<tr>
<td>733</td>
<td>-4</td>
<td>-22</td>
<td>-28</td>
<td>4.57</td>
<td>Left Brainstem</td>
</tr>
</tbody>
</table>
Figure 1.3. Whole-brain results.
Sympathy difference scores predict greater neural activation in altruists than controls for care-eliciting scenarios preceded by fearful expressions versus neutral expressions. Axial: $z = -23$ to $z = 51$, in increments of 6 mm. Sagittal: $x = -6$. Color bar = $t$ value. Clusterwise corrected threshold $p < .05$. Left = left.

We next aimed to evaluate functional coupling between the amygdala and periaqueductal gray in altruists and controls following depictions of vulnerability and distress. Functional coupling was assessed through generalized psychophysiological interaction (gPPI) analyses for which the anatomically defined left amygdala was the seed region of interest. The physiological variable was created through extraction of the de-convolved time series from this seed. gPPI analysis controls for functional connectivity with the seed region during other task conditions (including baseline), such that the resulting functional connectivity map is specific to the task condition of interest. A contrast test of functional connectivity with a periaqueductal gray region
of interest defined using the results of a recent meta-analysis (Linnman et al., 2012) compared functional coupling with the amygdala in altruists and controls as they reported their sympathy toward vulnerable protagonists following preattentively presented fearful expressions. Results revealed that, during this epoch, altruists exhibited increased functional connectivity between the left amygdala and left (but not right) periaqueductal gray ($x = -8, y = -28, z = -10, t(33) = 2.55, p < .05_{\text{corrected}}$; Figure 1.4). No cluster emerged in periaqueductal gray for which controls had greater functional connectivity with left amygdala than altruists.

![A and B: Functional connectivity results.](image)

Figure 1.4. Functional connectivity results. Altruists exhibit greater functional connectivity between left amygdala and left periaqueductal gray when rating sympathy for care-eliciting scenarios preceded by fearful expressions. (A) Left periaqueductal gray cluster. $z = -10$. Color bar $= t$ value. (B) Overlap between left periaqueductal gray cluster (red) and left periaqueductal gray ROI (yellow). $z = -10$. $k = 23$ ($2\text{mm}^3$ voxel size). Clusterwise corrected threshold $p < .05$.

Post hoc analyses examining movement across the task identified 1 altruist and 3 controls for whom at least 15% of the total number of task TRs were censored due to $> 0.5$ mm movement; all functional activation and connectivity analyses were repeated following exclusion of these participants; reported results were comparable following their removal (Table 1.4).
Table 1.4

Functional Activation and Connectivity Results in Altruists > Controls, Without Subjects Exceeding Motion Parameters

<table>
<thead>
<tr>
<th>k</th>
<th>Peak x</th>
<th>Peak y</th>
<th>Peak z</th>
<th>Peak t(28)</th>
<th>Peak Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Left Amygdala ROI Activation</td>
</tr>
<tr>
<td>11</td>
<td>-27</td>
<td>0</td>
<td>-26</td>
<td>3.14</td>
<td>Left Amygdala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Whole-Brain Functional Activation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>886</td>
</tr>
<tr>
<td>414</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Left Periaqueductal Gray ROI Functional Connectivity Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>40†</td>
</tr>
</tbody>
</table>

Note. Three controls and one altruist were excluded due to movement.  †2mm³ voxel size.

Structural connectivity

We then investigated whether these patterns corresponded to increased integrity of the white matter tracts between the amygdala and periaqueductal gray using probabilistic tractography of DTI data. Volumetric segmentation was performed on anatomical scans to obtain subject-specific anatomical seeds within the amygdala, which were then transformed into diffusion space. The periaqueductal gray target mask was hand-drawn for all subjects by a rater (K.O.) blind to subject group membership. Probabilistic tractography with multiple fiber orientations (Behrens et al., 2007) was run separately for left and right amygdala seeds. Then, given that left hemisphere fractional anisotropy (FA) values had unequal variance between groups (Levene’s test for homogeneity of variance, $F(1,36) = 13.41, p < .001$), non-parametric statistics were applied. Results revealed higher FA in the white matter tract connecting altruists’ left amygdala and periaqueductal gray relative to controls (altruist $M_{rank} = 23.89$, control $M_{rank} = 15.11; U = 264, p = .014$; Figure 1.5). In addition, FA in this tract correlated positively with self-reported empathy in altruists only, $r(15) = .61, p = .009$. This correlation was mainly
attributable to personal distress, \( r(15) = .61, p = .009 \) and empathic concern subscale scores, \( r(15) = .59, p = .013 \), the former of which survived Bonferroni correction for multiple comparisons. No comparable relationship between FA and empathy was found in controls. No group difference in FA was observed in the right amygdala-periaqueductal gray tract (altruist \( M \) rank = 17.89, control \( M \) rank = 21.11; \( U = 150, p = .385 \)).

![Figure 1.5](image)

*Figure 1.5. Structural connectivity in tracts linking amygdala and periaqueductal gray. (A) Left amygdala-periaqueductal gray tract. Orange = altruists, yellow = controls, red = overlap. (B) Altruists had higher mean FA values within the left amygdala-periaqueductal gray tract, \( p = .014 \). Error bars represent SEM.*

No group difference in the volume (in mm\(^3\)) of the tracts was observed for either the left tract (altruist \( M = 1,572, SD = 540 \); control \( M = 1,537, SD = 517 \); \( t(36) = 0.21, p = .838 \)) or right tract (altruist \( M = 1,532, SD = 470 \); control \( M = 1,478, SD = 483 \); \( t(36) = 0.35, p = .731 \)). When only including data from participants who had been retained in functional connectivity analysis (see Methods), the group FA difference in the left amygdala-periaqueductal gray tract persisted (altruist \( M \) rank = 21.59, control \( M \) rank = 13.41; \( U = 214, p = .016 \)) and, again, no comparable group difference was observed on the right (altruist \( M \) rank = 15.82, control \( M \) rank = 19.18; \( U = 116, p = .339 \)).
Discussion

Our results provide the most direct evidence to date that conserved parental care circuits support costly altruism toward strangers. Extraordinary altruists’ experienced sympathy following depictions of vulnerability and distress corresponded to heightened activation in both the left amygdala and periaqueductal gray; these regions also exhibited enhanced functional coupling in altruists relative to controls during this epoch, and were connected by higher-integrity white matter tracts, indicating that these regions may comprise a functional network that supports care-based decisions in extraordinary altruists. Supporting this interpretation, among altruists, individual variation in self-reported empathy corresponded to increases in the integrity of the white matter-tracts connecting amygdala and periaqueductal gray. Together, these findings, which identify consistent patterns across behaviorally-linked BOLD activity, functional connectivity, and structural connectivity analyses and which are consistent with existing literature regarding the basis of caregiving in non-human animal models, converge to support the importance of a subcortical caregiving system supported by amygdala-midbrain connections in the provision of costly altruism toward strangers. These findings add to accumulating evidence supporting distinct neural pathways underlying distinct forms of altruism. In contrast to the current findings, for example, reciprocal altruism relies primarily on frontal-striatal pathways that mediate expectations of immediate and future reward (Hein et al., 2016; Rilling & Sanfey, 2011; Rilling et al., 2002; Rilling et al., 2004) and kin-based altruism relies on limbic and cortical pathways that mediate responsiveness to familiar others (Lieberman et al., 2007; Platek & Kemp, 2009; Platek et al., 2005).

These findings may assist in understanding a seeming paradox of extraordinary altruism, which is that costly sacrifices for strangers by definition entail undertaking substantial risks (both
health-related and financial risks, in the case of kidney donation) and yet individuals who engage in this behavior exhibit heightened sensitivity to fear-relevant stimuli, including relatively better recognition of fearful facial expressions and heightened amygdala reactivity to these cues (Marsh et al., 2014). How can heightened fear sensitivity be associated with seemingly risk-insensitive behavior? Reciprocal connections between amygdala and periaqueductal gray may hold the answer. A critical feature of caregiving is the inhibition of behavioral avoidance via pathways that connect the central amygdala and periaqueductal gray (Viviani et al., 2011). Such a switch from avoidance to approach characterizes the onset of parental and alloparental caregiving across species (Preston, 2013). Responses to infantile distress cues have been studied most extensively in rodents, given the behavioral transformation that occurs during the transition to parenthood. Nulliparous females find infants aversive and actively avoid them, but hormonal cascades toward the end of gestation alter patterns of receptors for critical care-supporting neurotransmitters like oxytocin in the amygdala and periaqueductal gray, among other regions, causing cues associated with infantile vulnerability and distress to become highly appetitive (Insel, 1992; Mattson, Williams, Rosenblatt, & Morrell, 2001; Preston, 2013), and biasing the organism away from preservative behaviors and toward protective behaviors. Increases in maternal aggression to protect vulnerable rodent offspring are dependent on concurrent decreases in fear, and both phenomena are thought to be modulated by activity in periaqueductal gray (Lonstein & Gammie, 2002). In sum, the present data are consistent with evidence suggesting that amygdala-mediated detection of distress cues must be coupled with periaqueductal gray-mediated inhibition of preservative behaviors in order for protective caregiving responses to vulnerability and distress to emerge.
Our findings are also indirectly consistent with the previously postulated role for the neurotransmitters oxytocin and GABA in caregiving behaviors, including costly altruism. Both the amygdala and periaqueductal gray contain high concentrations of oxytocin (Jenkins, Ang, Hawthorn, Rossor, & Iversen, 1984) and oxytocinergic receptors (Vaccari, Lolait, & Ostrowski, 1998; Yoshimura et al., 1993), consistent with these structures’ role as part of an oxytocinergic social network (Keltner, Kogan, Piff, & Saturn, 2014). Exogenous administration of oxytocin has been shown to enhance protective behaviors in both humans and other mammalian species (Madden & Clutton-Brock, 2011; Mah, Bakermans-Kranenburg, Van IJzendoorn, & Smith, 2015; Pedersen, Ascher, Monroe, & Prange, 1982) and to modulate functional connectivity specifically between amygdala and a midbrain region proximal to the region identified here (x = -5, y = -25, z = -18) (Kirsch et al., 2005). GABA may also play an important modulatory role. GABAergic signaling in the ventrolateral periaqueductal gray, which is the location of highly-organized afferent amygdala GABAergic input (Tovote et al., 2016), also modulates offspring care behavior. For example, in rodents, GABA_A receptor antagonism in ventrolateral periaqueductal gray promotes licking and grooming of pups while decreasing anxiety and aggression (Lee & Gammie, 2010).

Observed results emerged despite limitations that included constrained sample sizes that reflect the extreme rarity of altruistic kidney donation. Also important to consider is that both amygdala and periaqueductal gray serve multiple sociobehavioral functions in addition to caregiving, including positive emotional experiences, attention, and physiological processes such as pain (Buhle et al., 2013; Linnman et al., 2012). However, our paradigm was designed to maximize the specificity of our conclusions. Our analytic approach for fMRI data identified only patterns of functional activation that corresponded to increases in reported sympathy during the
task, and the incorporation of a subject population of rare extraordinary altruists tethers our neuroimaging findings to objectively measured acts of costly altruism toward strangers outside the laboratory. Finally, the correspondence among our findings using multiple approaches, including functional and structural imaging, supports the significance of the identified regions in supporting altruistic caregiving. But, more precise understanding of the relevant neural processes will require more direct interrogation of the care system via genotyping or the use of neurochemical techniques such as ligand binding, which could strengthen conclusions about the precise molecular-level processes that support caregiving responses toward vulnerable non-kin others.

In sum, the current findings extend existing knowledge about the neural bases of costly altruism toward strangers. Altruistic kidney donation is a costly (Rodrigue et al., 2016), non-normative (Brethel-Haurwitz, Stoycos, Cardinale, Huebner, & Marsh, 2016) behavior performed to benefit an anonymous, non-kin other and thus meets the most stringent definitions of altruism (Batson, 2010; Clavien & Chapuisat, 2013; de Waal, 2008). In interviews, many altruistic donors report that the urge to donate their own kidney followed hearing or seeing a media story featuring a stranger suffering from kidney failure prior to receiving a donor kidney. Our findings are consistent with the possibility that these sympathetic urges in response to depictions of distressed and vulnerable others correspond to activation and structural integrity in a network of regions associated with prosocial and nurturing behaviors, including the amygdala and periaqueductal gray, both of which serve essential roles in motivating offspring care across mammalian species (Marsh, 2016; Nephew et al., 2009; Numan, 2006; Preston, 2013). Together, these results support the possibility that costly altruism toward strangers, rather than being a
wholly inexplicable outcome, may represent an exaptation of the mammalian parental care system.
STUDY 2: ALTRUISTS DEMONSTRATE ENHANCED SELF-OTHER OVERLAP IN PAIN PROCESSING IN ANTERIOR INSULA

Study 1 adds to evidence that highly altruistic individuals have a heightened sensitivity to distress and vulnerability in others (Marsh et al., 2014), and extends these prior findings by demonstrating that such sensitivity contributes to caring decision-making. However, what this neural sensitivity characterized by heightened amygdala responsiveness to distress means with regard to empathic responding remains to be determined. An extensive body of research exploring empathy for pain as a particular type of empathic responding suggests that neural simulation of pain underlies our empathic understanding of distress in others, in which affective pain regions such as the anterior insula and mid-cingulate cortex respond not only to the first-hand experience of pain but also to the observation of pain in another person (Bernhardt & Singer, 2012; Fan, Duncan, de Greck, & Northoff, 2011; Lamm et al., 2011; Zaki, Wager, Singer, Keysers, & Gazzola, 2016). By directly comparing the first-hand experience of pain with the observation of pain in others, the current study was designed to test if heightened neural sensitivity to distress in altruists represents enhanced neural simulation.

Extraordinarily altruistic individuals have demonstrated enhanced neurocognitive sensitivity to fear in others over two fMRI paradigms investigating sensitivity to distress and vulnerability. In an explicit passive viewing task, altruistic kidney donors were characterized by heightened right amygdala activation, specifically in response to fearful faces, and this amygdala activation predicted accuracy in recognizing such distress cues. Perhaps supporting this activation finding, altruists also had a larger right amygdala (Marsh et al., 2014). Such converging evidence from activation and structure also was found to support caring decision-making in Study 1. Caring decision-making in altruists predicted activation and functional connectivity between the amygdala and periaqueductal gray, two central structures in the
mammalian parental care system. These findings were corroborated by greater structural connectivity between these structures in altruists (see Study 1). Together, these results highlight a neurocognitive sensitivity to distress and vulnerability that appears to set altruists apart. But, what does this sensitivity suggest about empathic responding and altruism? Simulation accounts of empathy (de Vignemont & Singer, 2006; de Waal, 2008; Goldman & de Vignemont, 2009; Preston & De Waal, 2002), in which we understand the affective experiences of others through an empathic resonance in which affective states are at least partially simulated in ourselves, would suggest that these results should be interpreted as enhanced simulation in extraordinary altruists. However, thus far we have only examined altruists’ responses to distress in others, and not correspondence between observations of distress and their own first-hand experience of distress.

Extensive research exploring empathy for pain lends support to such a simulation hypothesis. Similar patterns of activation have been observed in anterior insula and anterior cingulate cortex during the first-hand experience and observation of pain in a loved one, an effect that was correlated with trait self-reported empathy. In contrast, activation in somatosensory cortex, posterior insula, and caudal anterior cingulate were observed only during the first-hand experience of pain (Singer et al., 2004). Such findings have since been corroborated by meta-analyses (Fan et al., 2011; Lamm et al., 2011) in which anterior insula and midcingulate cortex are consistently activated during empathy for pain across paradigms including the observation of pain cues, photos of painful situations, and empathizing with painful facial expressions. Such functional findings have recently been refined by multivoxel pattern analyses, which allow for more fine-grained spatial investigation of neural activation. These analyses have revealed both unique and shared patterns of activation in anterior insula and mid-cingulate cortex for self and
other experiences of pain (Corradi-Dell’Acqua, Hofstetter, & Vuilleumier, 2011; Corradi-Dell’Acqua, Tusche, Vuilleumier, & Singer, 2016). However, through their own multivoxel pattern analyses, Krishnan and colleagues (2016) argue that first-hand and vicarious pain are more distinct than previously described, and that the conclusions of other recent pattern analyses are limited in that shared patterns are not pain specific and appear to be more domain-general for shared negative affect, including disgust and fairness violations (Corradi-Dell’Acqua et al., 2016). Together, these findings highlight that it may be important to consider the aspects of the pain experience that are simulated, whether this be a representation of the sensory intensity of the pain (which was an important criterion for Krishnan and colleagues), or the negative affect associated with pain, which may be more relevant to empathy but be more general to experiences of distress (Zaki et al., 2016).

How an individual perceives their relationship with a person in distress, among other contextual factors (Hein & Singer, 2008), can be an important moderator of empathic simulation. While initial studies of empathy for pain capitalized on examining close relationships (e.g., romantic partners) (Singer et al., 2004), later investigations have highlighted that people are limited in who they may resonate with. Beckes and colleagues (2012) demonstrated that the correspondence in neural activation between experienced and observed threat of pain is greater for a friend, relative to a stranger, an effect that was further explained by self-reported closeness with the friend. Thus there may be limitations on whose emotions are simulated, affected by the closeness of our relationship with that person. Several paradigms have demonstrated that even relatively arbitrary relationships can modulate empathic simulations. Hein and colleagues (Hein et al., 2010) found that participants were more likely to help in-group members (sparing them from a painful shock by taking it themselves) than they were for out-group members. Further,
participants showed greater insula activation in response to the pain of an in-group member than an out-group member, and this insula activation bias was correlated with the helping difference between in-group and out-group members, all when group membership was defined by soccer team allegiance and all empathic targets were strangers. Given the extraordinary and costly act benefitting a stranger that characterizes altruistic kidney donors, and recent evidence that they show less of a tendency to discount outcomes for socially distant relations (Vekaria, Brethel-Haurwitz, Cardinale, Stoycos, & Marsh, under revision), altruists may demonstrate enhanced empathic simulation with strangers, responding to unknown others as the average individual tends to respond only to closer relations.

The current study was designed to interrogate self-other correspondence in pain processing, to test whether altruists’ sensitivity to strangers’ distress may be supported by empathic simulation. In an empathy for pain paradigm in which extraordinary altruists and matched controls experienced painful pressure to their right thumbnail and observed an unknown study partner during the same painful experience, it was hypothesized that altruists would demonstrate enhanced self-other overlap in affective pain regions including the anterior insula and mid-cingulate cortex that have been previously implicated in empathy for pain paradigms. Further, it was hypothesized that enhanced group-level self-other overlap would be supported by greater individual-level self-other correspondence in activation levels in these regions of interest. In other words, not only would altruists show more similar activation patterns for self and other pain, but also self activation levels would predict other activation levels across individual altruists. Such predictions are consistent with altruists’ demonstrated enhanced sensitivity to distress and vulnerability in others, their reduced tendency to discount outcomes for socially distant others, and their extraordinary costly altruism toward a stranger.
Methods

Participants

Study 2 was conducted 3 years after Study 1, and so constituted a separate experiment, albeit with partially overlapping subject samples, as 12 altruists and 6 controls returned for Study 2 procedures. Fifty-seven participants, who included 29 altruistic kidney donors and 28 matched controls, took part in this study for monetary payment. Due to scanner malfunction, one altruist was unable to complete the task. Additionally, three altruists and one control were excluded from analyses for movement exceeding 0.5 mm for at least 15% of the total number of task TRs. This resulted in a final sample of 25 altruistic kidney donors and 27 matched controls between 21 and 55 years old (see Table 2.1 for participant characteristics). Altruistic kidney donors were recruited using mailings and electronic advertisements through local and national transplant organizations. As in Study 1, the sample of altruists was limited by the extreme rarity of this behavior. Because altruists were recruited from across North America, most altruists resided more than a two-hour drive from the university and were provided with airfare and up to two nights’ lodging. All altruists had donated a kidney to a stranger unknown to them personally at the time of donation. Nineteen altruists were non-directed donors for whom the recipient was anonymous at the time of donation. The remaining six directed their donations to a specific individual who was known to them at the time of donation but whose need for a kidney they had learned about through, for example, a flier or Internet posting. All donations were verified through independent sources, including transplant center records or media reports. Healthy volunteers were recruited from the local community using fliers, online advertisements, and electronic participant databases including ResearchMatch.
Table 2.1

Study 2 Participant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Altruists (n = 25)</th>
<th>Controls (n = 27)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female (% Male)</td>
<td>9/16 (36.0%)</td>
<td>13/14 (48.2%)</td>
<td>.376</td>
</tr>
<tr>
<td>White/Other Race (% White)</td>
<td>23/2 (92.0%)</td>
<td>22/5 (81.5%)</td>
<td>.267</td>
</tr>
<tr>
<td>Right/Other Handed (% Right)</td>
<td>24/1 (96.0%)</td>
<td>25/2 (92.6%)</td>
<td>.599</td>
</tr>
<tr>
<td>Household Income ≥ $60,000</td>
<td>17 (73.9%)</td>
<td>24 (92.3%)</td>
<td>.109</td>
</tr>
<tr>
<td>Education ≥ Four-Year Degree</td>
<td>17 (68.0%)</td>
<td>23 (85.2%)</td>
<td>.142</td>
</tr>
<tr>
<td>Age M (SD)</td>
<td>41.92 (9.91)</td>
<td>38.89 (7.99)</td>
<td>.229</td>
</tr>
<tr>
<td>IQ M (SD)</td>
<td>108.28 (12.44)</td>
<td>111.22 (11.28)</td>
<td>.375</td>
</tr>
</tbody>
</table>

Note. One control did not report their household income.

As in Study 1, exclusion criteria for all participants included current use of psychotropic medication, history of head injury or neurological illness, IQ < 80 (as assessed using the Kaufman Brief Intelligence Test - Second Edition; Kaufman & Kaufman, 2004), and pregnancy or other contraindications to safe MRI scanning, including metal fragments or implants. In the current study, more thorough screening of psychopathology was also conducted, in order to ensure group matching on potentially relevant psychological variables. Participants were excluded if they scored above clinical cutoffs for Global Severity, Positive Symptom Distress, or Positive Symptom Total on the Symptom Checklist - 90 (SCL-90; Derogatis & Unger, 2010), particularly if these totals were driven by scores on the Somatization, Obsessive-Compulsive, Depression, Anxiety, Paranoid Ideation, or Psychoticism subscales. But, if these totals were driven by scores on the Interpersonal Sensitivity or Hostility subscales, participants were not excluded. Additionally, given the nature of the task, participants were excluded if they reported a pain disorder or hearing difficulties. Controls were excluded if they reported having ever volunteered to donate an organ to any individual (not including consenting to become a deceased organ donor), or if they expressed interest in potentially doing so. All study procedures were
approved by the Institutional Review Board at Georgetown University in Washington, DC, and all participants provided written informed consent before testing.

**Procedures**

All interested volunteers initially completed a 90-minute online survey assessing exclusion and inclusion criteria and demographic variables. Eligible volunteers were then screened by telephone to confirm eligibility. Researchers coordinated altruists’ travel to and lodging at Georgetown University to enable on-site neuroimaging and cognitive testing. To ensure groups were matched, eligible controls completed laboratory screening that included assessments of IQ, income, education, psychological history, medication use, and MRI compatibility before MRI scanning. After confirmation of eligibility, controls completed neurocognitive tasks and MRI scanning in a final visit.

During the calibration procedure and the fMRI task (see Neuroimaging Acquisition and Task), participants received painful pressure stimulation to their right thumbnail. Pneumatic pressure was administered remotely by an MRI-safe device connected to a laptop computer with software programmed to consistently administer a selected pressure level at intervals and durations matching the task structure. When pressure was administered, a circular plastic plunger with an 2 cm diameter rubber surface made contact with the right thumbnail of the participant or confederate, whose thumb was positioned in a clear plastic tube to keep it in place throughout the task (see Figure 2.1A). Pressure of a rectangular waveform was administered for 6 seconds.

Two live video feeds allowed each participant to observe both the hand of the confederate and their own hand during the task. A video camera positioned next to the confederate in the MRI console room captured video of the confederate’s hand. Each participant was briefly introduced to the confederate in this set-up, so that it was apparent to the participant that the
video feed would in fact be live. An MRI-safe video camera positioned adjacent to the scanner was adjusted for each participant to capture an equivalent view of the participant’s own right hand receiving painful stimulation. The video feed was switched between these two cameras depending on the run of the task.

**Neuroimaging Acquisition and Task**

MR images were acquired with a 3T Siemens Tim Trio scanner (Siemens Medical Solutions) and a 12-channel phased-array head coil. Functional data were collected using a T2*-weighted echo-planar imaging sequence (46 3.0 mm transversal slices; 64 × 64 matrix; repetition time, 2,500 ms; echo time, 30 ms; field of view, 192 mm²; 3.0 × 3.0 × 3.0 mm voxels). The first four volumes of each functional run were excluded from analysis to account for magnet stabilization. High-resolution T1-weighted anatomical images were also acquired (3D Magnetization Prepared Rapid Acquisition Gradient Echo; 176 1.0-mm axial slices; field of view, 250 mm²; repetition time, 1,900 ms; echo time, 2.52 ms; 246 × 256 matrix).

Participants completed three runs of the pain task, each lasting 12 minutes and 18 seconds. In the first two runs, participants observed a “study partner,” and in the third run they experienced the task themselves (Figure 2.1A). The study partner was a female confederate whom participants were briefly introduced to immediately prior to the scan. Two confederates participated over the course of the study, balanced between the two groups. The second run of the task began with a prompt to encourage increased empathizing with the study partner. Only the first and third runs will be compared in the current analyses, in which participants simply viewed the study partner complete the task and participants completed the task themselves, in order to examine group differences in baseline responses to observation of pain in another person relative to the first-hand experience of pain.
Figure 2.1. Trial structure of the empathic pain task.

(A) Participants viewed continuous live video feeds of the study partner’s hand in runs 1 and 2, then their own hand in run 3. (B) Variable anticipation periods were followed by the administration or omission of painful pressure stimulation to the right thumbnail. An audio cue at the beginning of the anticipation period indicated whether the trial was safe or had the potential for pain. A second audio cue indicated whether or not painful pressure was administered.

Each run had 30 trials, with each trial made up of a variable anticipation period, the administration or omission of painful pressure stimulation, and a variable rest period (Figure 2.1B). Half of the trials were safe trials, in which participants knew that there would be no thumb pressure, while the other half of the trials had the potential for thumb pressure. Pressure was omitted on one third of these potentially painful trials in order to keep the administration of pain probabilistic rather than deterministic, which is important to elicit fearful anticipation (Sylvers, Lilienfeld, & LaPrairie, 2011). Throughout each trial, participants viewed continuous live video feeds of either the confederate’s hand in the first and second runs, or their own hand in the third run. Audio cues via MRI-safe headphones indicated the trial structure, with a first sound indicating at the start of each trial whether it was a safe trial or had the potential for painful
pressure, and a second sound after the anticipation period indicating whether or not pressure was currently being administered. Participants were trained in the meaning of the various audio cues prior to the scan, reminded of the cue meanings once in the scanner, and quizzed on their comprehension prior to the first run. Participants’ ability to hear the audio cues and differentiate them from scanner noise was confirmed during a localizer scan prior to the first run.

Immediately prior to scanning, all participants were guided through a calibration procedure to determine the level of pressure to be administered during the scan, such that subjective pain perceptions were equivalent across participants. All participants were calibrated to receive pain that was subjectively “slightly intense” corresponding to a rating of 13.5 on the 21-point Gracely Box Scale, which was selected to be moderately painful but not harmful. Following explanation of the Gracely Box Scale and the calibration procedure, increasing pressure levels were administered, beginning with 5 psi and increasing by increments of 5 psi until participants gave a rating of 13.5. Ascending pressure was continued until a rating higher than 13.5 was given, then descending pressure in increments of 5 psi was administered until participants gave a rating lower than 13.5. If different, the lower of the two pressure levels rated 13.5 was selected as the “slightly intense” pressure level. The pressure selected during the calibration procedure was checked in the scanner immediately prior to the run in which participants received pressure themselves. Several participants in each group (8 altruists and 10 controls, $\chi^2(1) = 0.15, p = .703$) rated the pressure selected during calibration more highly than 13.5 once in the scanner, resulting in a re-calibration procedure in which descending pressure in increments of 5 psi was tested until participants gave a rating of 13.5. This re-calibrated lower pressure was then the pressure administered to the participant during the scan. Since altruists selected significantly higher objective pressure for their subjectively calibrated “slightly
intense,” $t(50) = 3.97$, $p < .001$ (Figure 2.2), final pressure level was included as a covariate in analyses. Importantly, subjective ratings during and after the scan confirmed that final pressure levels were subjectively equivalent for the two groups (see Results). The confederate always received 15 psi of pressure, which was selected to be clearly visible over the live video feed and within the range of slightly intense pain for both groups (see Figure 2.2).

![Figure 2.2](image)

Figure 2.2. Final subjective “slightly intense” pain level selected by each group. Contours represent frequency distributions. Error bars represent 95% confidence intervals based on the standard error of the mean, dots represent means.

Subjective ratings of pain experience were collected during and after scanning. Following each run, participants were asked to rate pain intensity on a 1 (no pain) to 7 (extreme pain) scale. For runs 1 and 2, this was their perception of the study partner’s pain level, while for run 3 this was their rating of their own pain experience. After the scan, participants filled out a questionnaire in which they rated how fearful and unpleasant the anticipation and pain portions of the task were, respectively, on a 1 (not at all) to 7 (extremely) scale. They also answered the same questions regarding the experience of their study partner. Participants also rated their
perception of closeness to others on the Inclusion of Other in the Self Scale (Aron, Aron, & Smollan, 1992). Closeness was rated by selecting one of seven diagrams of overlapping circles representing perceived closeness to the other individual, from 1 (separate and distant circles, representing lowest degree of closeness) to 7 (completely overlapping circles, representing highest degree of closeness). Relationships rated included immediate family, extended family, friend, acquaintance, stranger, and study partner from the empathic pain task. Data from one control were missing for the experience questionnaire and one altruist did not complete the Inclusion of Other in the Self Scale.

Analysis of Neuroimaging Data

Functional data were preprocessed and analyzed according to the general linear model, using Analysis of Functional NeuroImages (AFNI; Cox, 1996). The three runs of the task were concatenated, despiked, motion-corrected, and spatially smoothed using a 6 mm full-width half maximum Gaussian filter. Functional data were aligned to the anatomical grid, transformed to the Talairach and Tournoux Atlas (Talairach & Tournoux, 1988), and masked with an extents mask to account for motion artifacts and to exclude voxels without valid data at every TR for every run, helping to control for false activations. Twelve regressors were created to model the anticipation and stimulation portions of each run, resulting in the following four regressors for each run: fear (cue indicates potential pain), no fear (safety cue), pain (stimulation), no pain (no stimulation). For each participant, contrasts were also calculated for pain (pain > no pain) and fear (fear > no fear) for each run. Rest was modeled implicitly; baseline was modeled by a first-order function, and motion artifacts were modeled using the six estimated rigid-body motion parameters. Boxcar regressors representing the occurrence of each block type were convolved with a canonical hemodynamic response function, scaled to an amplitude of 1.
Group-level analyses were limited to a brain mask defined by voxels with functional activation shared by at least 50% of participants. Cluster size thresholds were calculated for a corrected clusterwise \( p \) threshold of .05 for this group mask (whole brain clusterwise corrected \( p = .05 \); 84 contiguous voxels at uncorrected \( p = .01 \)) using 10,000 Monte Carlo simulations conducted via 3dClustSim in AFNI (June 2016 AFNI 16.1.26 version of 3dClustSim, utilizing a Gaussian plus mono-exponential spatial autocorrelation function to decrease risk of false positives). MNI coordinates of peak \( t \) statistics within significant clusters are reported. As noted above, final administered pressure stimulation level was included as a centered covariate in all analyses given the significant group difference, to examine group differences due to subjective experience rather than objective pressure stimulation level. The covariate was centered across all participants, so as to not artificially remove group differences in this variable.

In order to examine group differences in similarity of neural activation for \textit{self} and \textit{other} experiences of fear and pain, conservative conjunction-null analyses (Nichols, Brett, Andersson, Wager, & Poline, 2005) were conducted in which activation maps meeting statistical thresholds in each condition were compared, such that regions of conjunction indicated significant activation in both conditions independently, an approach that has been utilized in examining similarity between \textit{self} and \textit{other} responses in empathy for pain paradigms (e.g., Beckes et al., 2012). Specifically, a mask was created from a thresholded \textit{self pain} > \textit{self no pain} contrast, and this mask was then applied to an \textit{other pain} > \textit{other no pain} contrast, within each group, and then group differences in self-other conjunctions were examined. Following up on conjunction patterns observed at the group level, covariation analyses were conducted, in which the association between \textit{self} and \textit{other} activation levels in regions of interest were tested, to determine if apparent self-other mapping at the group level in conjunction analyses were
supported by individual-level correspondence in neural activation. Regions of interest (ROIs) in bilateral anterior insula and midcingulate cortex were defined by spheres of 10 mm diameter centered at coordinates (left anterior insula: x = -40, y = 22, z = 0; right anterior insula: x = 39, y = 23, z = -4; midcingulate cortex: x = -2, y = 23, z = 40) derived from a functional meta-analysis of empathy for pain studies (Lamm et al., 2011). Cluster size thresholds were also calculated for a corrected clusterwise $p$ threshold of .05 for each of these ROIs (ROI clusterwise corrected $p = .05$: 2 contiguous voxels at uncorrected $p = .01$ in each ROI).

**Results**

**Neural Simulation of Pain**

Comparing activation for the first-hand experience of pain with activation for the observation of pain in a stranger, altruists exhibit greater self-other overlap in bilateral anterior insula, meaning that activation in bilateral anterior insula is significant for both the self and other conditions during pain, $p < .05_{\text{corrected}}$. While both altruists and controls had overlap in bilateral inferior parietal cortex and right anterior insula, only altruists exhibited significant overlap in left anterior insula (Figure 2.3, Table 2.2).
Figure 2.3. Other pain results.

Altruists exhibit enhanced bilateral anterior insula activation for pain other > no pain other, masked with pain self > no pain self. Anterior insula ROI spheres are displayed (white circles) in the middle image for each group. Color bar = t value. Clusterwise corrected threshold $p < .05$.

Table 2.2

Regions of Significant Activation for Other Pain > Other No Pain, Masked by Self Pain > Self No Pain, in Each Group

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th>Peak x</th>
<th>Peak y</th>
<th>Peak z</th>
<th>Peak $t$ (23)</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Altruists</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>588</td>
<td>59</td>
<td>18</td>
<td>7</td>
<td>8.33</td>
<td>Right Inferior Frontal Gyrus, Right Insula</td>
</tr>
<tr>
<td></td>
<td>554</td>
<td>66</td>
<td>-38</td>
<td>24</td>
<td>9.39</td>
<td>Right Inferior Parietal Lobule</td>
</tr>
<tr>
<td></td>
<td>548</td>
<td>50</td>
<td>-59</td>
<td>-5</td>
<td>9.23</td>
<td>Right Inferior Temporal Gyrus</td>
</tr>
<tr>
<td></td>
<td>511</td>
<td>-62</td>
<td>-30</td>
<td>41</td>
<td>8.25</td>
<td>Left Inferior Parietal Lobule</td>
</tr>
<tr>
<td></td>
<td>455</td>
<td>-49</td>
<td>-74</td>
<td>-4</td>
<td>10.44</td>
<td>Left Middle Occipital Gyrus</td>
</tr>
<tr>
<td></td>
<td>213</td>
<td>-49</td>
<td>13</td>
<td>24</td>
<td>5.65</td>
<td>Left Inferior Frontal Gyrus, Left Insula</td>
</tr>
<tr>
<td></td>
<td>98</td>
<td>-30</td>
<td>-62</td>
<td>-55</td>
<td>5.29</td>
<td>Left Cerebellum</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2474</td>
<td>50</td>
<td>-68</td>
<td>-8</td>
<td>10.74</td>
<td>Right Middle Occipital Gyrus</td>
</tr>
<tr>
<td></td>
<td>1407</td>
<td>-40</td>
<td>-83</td>
<td>7</td>
<td>10.75</td>
<td>Left Middle Occipital Gyrus</td>
</tr>
<tr>
<td></td>
<td>925</td>
<td>-56</td>
<td>-27</td>
<td>41</td>
<td>9.43</td>
<td>Left Postcentral Gyrus, Left Inferior Parietal Lobule</td>
</tr>
<tr>
<td></td>
<td>478</td>
<td>53</td>
<td>11</td>
<td>35</td>
<td>10.42</td>
<td>Right Inferior Frontal Gyrus, Right Insula</td>
</tr>
</tbody>
</table>
These differences emerged despite similarity in thresholded whole-brain self pain > self no pain masks. No significant differences emerged between groups in a whole-brain contrast. Both altruists and controls had increased activation in bilateral insula, cingulate cortex, somatosensory cortex, inferior parietal cortex, dorsolateral prefrontal cortex, periaqueductal gray, striatum, and thalamus (Figure 2.4, Table 2.3).

Altruists

Controls

Figure 2.4. Self pain results.
Altruists and controls exhibit similar neural activation for self pain > self no pain. Anterior insula ROI spheres are displayed (white circles) in the middle image for each group. Color bar = t value. Clusterwise corrected threshold p < .05.
Table 2.3

Regions of Significant Activation for Self Pain > Self No Pain in Each Group

<table>
<thead>
<tr>
<th></th>
<th>Altruists</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k</td>
<td>Peak x</td>
</tr>
<tr>
<td>5325</td>
<td>-40</td>
<td>-2</td>
</tr>
<tr>
<td>1254</td>
<td>28</td>
<td>-51</td>
</tr>
<tr>
<td>803</td>
<td>-30</td>
<td>-63</td>
</tr>
<tr>
<td>792</td>
<td>-1</td>
<td>8</td>
</tr>
<tr>
<td>515</td>
<td>-4</td>
<td>25</td>
</tr>
<tr>
<td>475</td>
<td>37</td>
<td>-20</td>
</tr>
<tr>
<td>164</td>
<td>12</td>
<td>-49</td>
</tr>
<tr>
<td>144</td>
<td>-30</td>
<td>-14</td>
</tr>
<tr>
<td>132</td>
<td>-40</td>
<td>-84</td>
</tr>
<tr>
<td>101</td>
<td>34</td>
<td>-17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k</td>
<td>Peak x</td>
</tr>
<tr>
<td>6289</td>
<td>56</td>
<td>-26</td>
</tr>
<tr>
<td>5758</td>
<td>-53</td>
<td>-24</td>
</tr>
<tr>
<td>223</td>
<td>8</td>
<td>-32</td>
</tr>
<tr>
<td>175</td>
<td>56</td>
<td>45</td>
</tr>
<tr>
<td>131</td>
<td>-33</td>
<td>43</td>
</tr>
</tbody>
</table>

These group differences in self-other conjunction were confirmed with ROI analyses in bilateral anterior insula and midcingulate cortex – regions consistently activated for both self and other pain across prior studies (Lamm et al., 2011). Altruists, but not controls, had significant self-other overlap in activation in left anterior insula (altruists: k = 5, t(23) = 4.13, p < .05<sub>corrected</sub>). Both altruists and controls had significant overlap in pain processing in right anterior insula (altruists: k = 19, t(23) = 5.18, p < .05<sub>corrected</sub>; controls: k = 12, t(25) = 4.05, p < .05<sub>corrected</sub>), but neither group had significant overlap in midcingulate cortex.

Mean parameter estimates were extracted for regions of interest in bilateral anterior insula to examine individual-level correlation in activation levels for the self and other pain processing. Partial correlations controlling for objective pain level confirmed group-level findings in left insula. In altruists, activation in left anterior insula during the first-hand
experience of pain predicted activation in this same region in response to pain experienced by a stranger, $r(22) = .57, p = .004$, while no such relationship was observed in controls, $r(24) = -.15, p = .456$ (Figure 2.5). Neither altruists, $r(22) = .19, p = .380$, nor controls, $r(24) = .22, p = .291$, exhibited significant self-other covariation in right anterior insula.

![Figure 2.5. Self-other correlation in BOLD activation in left anterior insula ROI.](image)

Further, there was a trend toward less of a difference between activation levels in this left anterior insula ROI between the *self* and *other* conditions for altruists relative to controls, in which the absolute value of differences in activation levels between *self* and *other* across participants in each group was lower for altruists than controls, $t(50) = 1.77, p = .083$. Thus, altruists demonstrate more significant overlap in left anterior insula in *self* and *other* pain processing at the group-level in a conjunction analysis, individual-level covariation in activation levels suggests this overlap represents similar self-other responses, and altruists also a trend
toward greater self-other correspondence in how active the left anterior insula is in response to both types of pain.

**Inclusion of Other in the Self**

Greater self-other correspondence in pain processing may have been due to increased perceptions of closeness with the study partner in altruists, which was confirmed by ratings on the Inclusion of Other in the Self Scale. Altruists rated greater closeness with the study partner, $t(48) = 2.16$, $p = .036$, and there was a trend toward altruists rating greater closeness with strangers in general, $t(48) = 2.00$, $p = .051$. As expected, based on prior findings that extraordinary altruists do not differ in their perceptions or valuations of objectively closer others (Vekaria, Brethel-Haurwitz, Cardinale, Stoycos, & Marsh, under revision), altruists and controls did not differ in perceptions of closeness with objectively closer relationships, all $p > .05$ (Figure 2.6). Closeness with the study partner did not directly predict other neural activation or subjective ratings in either group.

*Figure 2.6. Perceptions of closeness as rated on the Inclusion of Other in the Self Scale. Error bars represent 95% confidence intervals based on the standard error of the mean. *$p < .05$. 
Subjective Experience

Despite significantly different objective pressure stimulation levels selected as “slightly intense” by altruists and controls (Figure 2.2), subjective experiences of fear and pain were equivalent between the two groups. There were no group differences in subjective ratings of fear or pain, during or after the scan, for the participant’s first-hand experience or their perception of the study partner’s experience, all $p > .05$ (Table 2.4). Across all participants, cues indicating that pain may occur were rated as more fear-inducing than safety cues, $t(51) = 12.44, p < .001$, and thumbnail stimulation was rated as more unpleasant than the omission of thumbnail stimulation, $t(51) = 12.90, p < .001$. Threat of pain, $t(51) = 13.91, p < .001$, and the experience of thumbnail stimulation, $t(51) = 11.92, p < .001$, were also rated negatively for the study partner, relative to the safety trials and omission of stimulation.

Table 2.4

<table>
<thead>
<tr>
<th>Rating</th>
<th>Altruists</th>
<th>Controls</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self Pain</td>
<td>4.48 (1.23)</td>
<td>4.30 (1.30)</td>
<td>.603</td>
</tr>
<tr>
<td>Other Pain</td>
<td>4.08 (1.08)</td>
<td>3.93 (0.83)</td>
<td>.564</td>
</tr>
<tr>
<td>Self Pain Unpleasantness</td>
<td>3.60 (1.38)</td>
<td>3.85 (1.38)</td>
<td>.527</td>
</tr>
<tr>
<td>Other Pain Unpleasantness</td>
<td>3.84 (1.28)</td>
<td>3.81 (1.58)</td>
<td>.936</td>
</tr>
<tr>
<td>Self Fear</td>
<td>3.36 (1.52)</td>
<td>3.65 (1.16)</td>
<td>.442</td>
</tr>
<tr>
<td>Other Fear</td>
<td>3.48 (1.39)</td>
<td>3.92 (1.16)</td>
<td>.222</td>
</tr>
</tbody>
</table>

Note. Mean (SD). During the scan, participants rated subjective pain for themselves and the study partner on a 1 (no pain) to 7 (extreme pain) scale. After the scan, participants rated how fearful and unpleasant the anticipation and pain portions of the task were, respectively, on a 1 (not at all) to 7 (extremely) scale, for both themselves and the study partner. This pattern of results is the same when controlling for objective pain level.

Associations between objective pain, self-reported subjective pain and unpleasantness, and neural activation in ROIs across all participants were considered next. While objective pain predicted right anterior insula and midcingulate activation during the first-hand experience of
pain, subjective pain only predicted midcingulate activation, and perceived unpleasantness did not predict mean neural activation in any of the ROIs (Table 2.5). Multiple regression analyses including both objective and subjective pain levels as predictors revealed that objective pain better predicted activation in both the right anterior insula, $\beta = .42, t(49) = 3.05, p = .004$, and midcingulate, $\beta = .32, t(49) = 2.38, p = .021$. In bivariate correlations across all subjects, subjective ratings of other pain and unpleasantness did not predict activation in ROIs during the observation of pain in the study partner, all $p > .05$.

Table 2.5

<table>
<thead>
<tr>
<th></th>
<th>Final PSI</th>
<th>Pain</th>
<th>Unpleasantness</th>
<th>Left AI</th>
<th>Right AI</th>
<th>Midcingulate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final PSI</td>
<td>-</td>
<td>.30*</td>
<td>.00</td>
<td>.18</td>
<td>.39**</td>
<td>.38**</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unpleasantness</td>
<td></td>
<td></td>
<td></td>
<td>.43**</td>
<td>.23</td>
<td>.03</td>
</tr>
<tr>
<td>Left AI</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>.03</td>
<td>.09</td>
</tr>
<tr>
<td>Right AI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.30*</td>
</tr>
<tr>
<td>Midcingulate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. AI = anterior insula.  
*p < .05. **p < .01.

Discussion

The current study extends prior findings supporting a role for neural simulation in empathy and altruism, suggesting that empathic resonance is enhanced in extraordinarily altruistic individuals. In a closely matched empathic pain paradigm, altruistic kidney donors were characterized by greater self-other overlap in pain processing in bilateral anterior insula, a region that has been implicated in both the first-hand and vicarious experience of pain across numerous studies. Further, altruists had greater correspondence between self and other activation levels in left anterior insula, highlighting that group-level overlap in this region was supported by the
individual-level prediction of *other* pain activation by *self* pain activation. *Self* and *other* activation levels in left anterior insula also exhibited a trend toward greater similarity in degree of activation, in that there was a greater match between *self* and *other* activation levels in this region of interest for altruists. Together, these findings suggest that when altruists are observing a stranger in pain, neural processing in anterior insula is more closely matching the representation of their own experience of pain. According to many dominant models of empathy, such neural resonance is an important empathic process for driving any subsequent helping behavior in observing someone else in pain or distress.

This empathic simulation interpretation is based in theories of embodied social cognition. Theories such as the Perception Action Model (de Waal, 2012; Preston & De Waal, 2002), in which understanding of and concern for another’s physical or emotional experience results from the activation of aspects of such experiences in ourselves, are consistent with extensive evidence that empathic pain is characterized by activation in affective components of the neural pain matrix also activated for the first-hand experience of somatic pain, including bilateral anterior insula (Lamm et al., 2011). More recent multivoxel pattern analyses have built on prior functional findings, highlighting that similar patterns in anterior insula respond to both *self* and *other* pain (Corradi-Dell’Acqua et al., 2011). The pattern associated with pain also characterized responses to other negative stimuli (Corradi-Dell’Acqua et al., 2011) and the anterior insula has been implicated in affective empathy more generally (Fan et al., 2011; Kanske, Bockler, Trautwein, & Singer, 2015; Tusche, Bockler, Kanske, Trautwein, & Singer, 2016). The anterior portion of the insula in particular is thought to be an affective center sensitive to perceived unpleasantness (Craig, Chen, Bandy, & Reiman, 2000; Rainville, Duncan, Price, Carrier, & Bushnell, 1997) that integrates somatic and interoceptive information from middle and posterior
insula (Craig, 2003; Craig, 2009), giving this region a role in understanding not only first-hand feelings and body-states, but also those of others (Lamm & Singer, 2010). Further, the anterior insula is hypothesized to be critical for learning and predicting emotional states in social settings (Singer, Critchley, & Preuschoff, 2009). While the group difference in objective pressure level rated as subjectively “slightly intense” was treated as a covariate of no interest in the current analyses, such a finding could have interesting implications for interoceptive differences between the two groups. That altruists seem to have a higher pain threshold and also show greater matching between self and other pain processing in anterior insula suggests that the integrative interoceptive processing of this neural region may be distinct in highly altruistic individuals. Analyses examining functional connectivity of the anterior insula during self and other pain processing may further clarify interoceptive differences.

Perhaps consistent with these various roles of the insula in emotional and empathic processes, other recent evidence suggests that the neural signatures for first-hand and empathic pain experiences are distinct, specifically that multivoxel patterns tracking the intensity of each type of experience are dissociable. Krishnan and colleagues (Krishnan et al., 2016) provide evidence for distinct patterns of neural activation for first-hand and vicarious experiences of pain, which are not mutually predictive. Further, they highlight that the role of the anterior insula in responding to distress and negative affective experiences more generally, particularly findings that neural patterns responsive to vicarious pain also respond to non-painful negative images (Corradi-Dell’Acqua et al., 2011) and vicarious disgust and unfairness (Corradi-Dell’Acqua et al., 2016), suggest limitations on simulation theories. Their findings raise important questions about the specificity of neural correspondence observed in studies to date and the implications of non-specificity for claims of empathic resonance. But, as suggested by Zaki and colleagues (Zaki
et al., 2016), while patterns of activation tracking objective pain intensity (such as those examined by Krishnan and colleagues) are useful in delineating similarities and differences between somatic and vicarious pain, this may reflect an aspect of the pain experience distinct from subjective affective components that play an important role in empathy. Whether the self-other overlap observed here is specific to the pain experience, or more general to distress and negative affect, remains to be determined. However, whether the effect was specific or general, it appears to be enhanced in altruists, who are known to also be sensitive to fear in others (Marsh et al., 2014, Study 1), and may be characterized by an enhanced domain-general distress sensitivity.

A strength of the current study is the close matching in the first-hand experience of pain and the pain experience observed in a stranger. Both conditions featured nearly identical live video feeds of the right hand of either the participant or study partner experiencing the same painful pressure stimulation, within the same paradigm. Other studies have relied on secondary cues (e.g., Corradi-Dell’Acqua et al., 2016; Hein et al., 2010; Singer et al., 2004; Singer et al., 2006) to indicate whether pain was occurring, or photos of painful situations distinct from the pain experienced by the participant (e.g., Krishnan et al., 2016; Lamm, Meltzoff, & Decety, 2010; Lamm, Nusbaum, Meltzoff, & Decety, 2007). While the current study focused on carefully matched subjective experiences, future investigations of empathic simulation in highly altruistic individuals could explore the effects of somatic and vicarious pain experiences of varying intensities, which could help to determine if the patterns observed here are specific to pain or more generally apply to distress (Krishnan et al., 2016; Zaki et al., 2016). Further, while the current study found greater self-other overlap in left anterior insula in altruists, which was supported by a positive association between self and other activation levels at the individual-level, such univariate analyses cannot determine if functional overlap reflects common
computational mechanisms as does multivoxel pattern analyses applied to empathic processing. While the current results suggest a special role for the anterior insula in processing both self and other pain in altruists, it does not determine how the insula is contributing to each of these processes, and cannot specify if it is single group of neurons that is responding to both self and other pain, or two groups of distinct but spatially contiguous neurons responding specifically to each experience. Applying multivoxel pattern analyses to the current dataset, and also comparing neural activation for self and other experiences of fearful anticipation to self and other experiences of pain, could further elucidate similarities and differences between these empathic processes.

Of note, it has been argued that vicarious distress and self-other overlap can be problematic for the altruistic nature of helping. Specifically, Cialdini and colleagues (Cialdini & Kenrick, 1976; Cialdini et al., 1987; Cialdini, Brown, Lewis, Luce, & Neuberg, 1997) have raised the concern that vicarious distress is a self-focused rather than other-regarding emotional response, and any resulting helping behavior is motivated by self-interested regulation of the vicarious distress, rather than relieving the suffering of the person helped. Several aspects of the current study suggest that altruists were not more susceptible than controls to a personal distress reaction. While altruists demonstrated greater self-other overlap in left anterior insula, this effect was not driven by significantly greater activation in altruists than controls in either condition in this region of interest. Thus, it was not due to more of either self or other pain-related activation in this region that altruists demonstrated greater self-other correspondence than controls. Subjective ratings for self and other pain also did not differ between the two groups.

Results of the current study suggest that enhanced empathic resonance may set highly altruistic individuals apart. These findings extend other recent findings on altruists’ sensitivity to
distress and vulnerability (Marsh et al., 2014; Study 1) and decreased tendency to discount outcomes for socially distant others (Vekaria, Brethel-Haurwitz, Cardinale, Stoycos, & Marsh, under revision). Specifically, the similarities between self and other pain processing observed here, in which extraordinary altruists may be responding to the pain of a stranger as most people respond to the pain of someone closer to them, support hypotheses that empathic simulation may underlie altruists’ responsiveness to suffering or distress in strangers and willingness to engage in costly and risky helping behavior on their behalf.
STUDY 3: EMPATHIC EMOTION REGULATION

Studies 1 and 2 have highlighted that low-level emotional processes can be automatically invoked in response to distressed and vulnerable others, particularly in highly altruistic individuals. As Vaish reviewed recently (2016), it is thought that our concerned responses to others are a result of not only these seemingly bottom-up emotional responses, but also top-down cognitive processes that can modulate the behavioral effect of such emotional responses, resulting in a “multidetermined concern.” The influence of modulatory processes on emotions evoked in response to another’s distress, particularly emotion regulatory processes, may be important for biasing an individual toward a concerned and compassionate response rather than personal distress (Batson et al., 1987; Batson, O’Quin, Fultz, Vanderplas, & Isen, 1983; Eisenberg et al., 1989a; Klimecki, Leiberg, Ricard, & Singer, 2014). However, another threat to a concerned prosocial response is regulation of empathic affect resulting in apathy (Cameron & Payne, 2011). The current studies thus sought to examine the role of emotion regulatory appraisals in promoting prosocial helping behavior, first in a proof of concept test of an empathic emotion regulation task, and then in an extension of this concept to extraordinary altruists.

Emotion regulation is the set of processes by which we can change the occurrence, strength, and duration of negative or positive emotional reactions (Gross, 2015). Specifically, it is the enactment of a goal to modulate an emotion when that emotion is inconsistent with a desired state. Just as effective emotion regulation helps to keep intrapersonal emotional experiences in check, it may help shape empathic emotion into a concerned response to distress that will result in prosocial helping. Empathic reactions can be dominated by either self-focused personal distress or other-focused empathic concern (Batson et al., 1987; Eisenberg, 2000). While empathic concern predicts helping behavior, personal distress predicts greater negative
arousal and avoidance of helping (Batson et al., 1987; Carrera et al., 2013; Eisenberg & Miller, 1987; Eisenberg et al., 1994; Eisenberg, 2000; Lopez-Perez, Carrera, Ambrona, & Oceja, 2014). Decety (2010) has argued that along with emotion understanding and empathic arousal, emotion regulation is a core component of empathic processing that may translate to prosocial helping. Similarly, Eisenberg and colleagues (Eisenberg & Okun, 1996; Eisenberg et al., 1994; Eisenberg et al., 1996; Eisenberg et al., 1998; Eisenberg, 2000; Fabes et al., 1994) have suggested that the development of emotion regulation is critical for the development of empathy and prosocial tendencies. In particular, they hypothesized that while personal distress and sympathy (their term for empathic concern) would both be predicted by emotional reactivity, emotion regulation would positively predict sympathy while negatively predicting personal distress (Eisenberg et al., 1994). They confirmed these hypotheses empirically in adults, finding that dispositional high reactivity and low regulation predicted personal distress, while moderate reactivity and regulation predicted sympathy (Eisenberg et al., 1994). Thus, when people experience negative affect as a result of another’s distress, they may activate a goal to regulate this level of negative affect to within a range that would bias them toward helping rather than avoiding the needy other.

Cognitive reappraisal, an emotion regulatory strategy in which one changes their thinking about an emotional situation in order to change their reaction to it, is one of the most studied forms of emotion regulation, partly because it is thought to be one of the most effective (Gross, 1998a). Cognitive reappraisal in particular has been found to cause both increased empathic concern (Lopez-Perez & Ambrona, 2014) and decreased personal distress associated with decreased amygdala and insula activation while observing someone else’s pain (Lamm, Batson, & Decety, 2007). In one of the few studies thus far to link emotion regulation and empathy to
situational prosocial outcomes, the tendency to engage in cognitive reappraisal predicted empathic concern and helping behavior toward an individual in distress, while expressive suppression (another form of emotion regulation characterized by the inhibition of outward displays of emotion) was negatively associated with prosociality (Lebowitz & Dovidio, 2015). Thus, with regard to prosocial outcomes, not all forms of emotion regulation may be effective in the regulation of vicarious negative affect. If personal distress is a result of overwhelming resonant negative affect in response to another’s distress, growing evidence suggests that cognitive reappraisal techniques that can be used to reduce feelings of fear or anxiety in an intrapersonal setting can also be applied to interpersonal emotions.

The role of emotion regulation in promoting prosociality is not without limitations, however. Cameron and colleagues (2011) have highlighted that emotion regulation can also inhibit empathy and prosociality, and their “collapse of compassion” studies have found that higher trait emotion regulation skill and also instructed emotion regulation both resulted in decreased concern for multiple victims. Similarly, Lockwood and colleagues (2014) have found that trait empathy only predicts prosocial tendencies for those with low to moderate trait reappraisal tendency. Such studies highlight that not all emotion regulation strategies may be equivalent with regard to interpersonal outcomes, and the goals and context of emotion regulation may be critical. Specifically, whether the goals of empathic emotion regulation are other-regarding or self-focused may be a key consideration. This limitation was considered in the design of the empathic emotion regulation task implemented in the current set of studies. While a common strategy in many studies examining intrapersonal emotion regulation via cognitive reappraisal is to decrease negative affect via a distancing strategy (i.e., imagining the situation as fake or otherwise distant from the immediate reality of the observer), such a strategy would not
be expected to be compatible with the goal of empathizing with a distressed victim or ultimately engaging in behavior that may help them. Thus, the goal of regulated emotions and the type of reappraisal utilized are additionally important in determining if the resulting motivation will be self-focused or other-focused.

Keeping with these “collapse of compassion” findings, recent research has suggested that much like a quadratic relationship between stress and performance on a variety of tasks known as the Yerkes-Dodson Law (Yerkes & Dodson, 1908), in which too little or too much stress is detrimental, there may be a similar inverted U-shaped relationship between simulated negative affect and prosocial tendencies. This possibility is consistent with the quadratic relationship observed between vagal activity and prosociality observed by Kogan and colleagues (Kogan et al., 2014). Vagal tone, measured via respiratory sinus arrhythmia as an index of parasympathetic control over heart rate, was quadratically related to prosocial traits, emotions, and third-party ratings, such that a moderate amount of vagal tone was most strongly associated with prosociality, while this was less true for low and high vagal tone. They hypothesized that because increased activity of the vagus nerve is associated with decreased heart rate (also perhaps greater tendencies toward socialization), it may play a role in the regulation of personal distress, which would contribute to prosocial behavior. Too much vagal activity, and thus greater deceleration of heart rate, is thought to be analogous to over-regulated negative affect, in the sense that there is such a decrease in negative affect that the motivator of empathic distress has been eliminated. Such an association between negative affect and prosociality may be conserved across species. High levels of the stress hormone corticosterone and also the administration of an anxiolytic drug decreased helping tendency toward a trapped conspecific in rats (Ben-Ami Bartal et al., 2016), meaning that both high stress and dampening of anxiety caused decreases in helping
behavior. Together, these results again suggest that there may be an optimal range of negative affect with regard to prosociality, produced by sufficient empathic resonance of negative affect paired with emotion regulation tendency such that this empathic affect does not become aversive and self-focused.

The present research was conducted in order to explore how affect and reappraisal may interact to influence prosociality toward distressed and vulnerable others. Given the mixed evidence cited thus far, two reappraisal strategies with potentially opposing effects on prosocial behavior were tested in the context of a donation task. A reappraisal strategy expected to be effective was a hopeful reappraisal in which participants attempted to decrease negative empathic affect, but reframe the situation in a particularly optimistic way. Given the approach-oriented nature of this reappraisal, and thus its similarity to a compassionate orientation (Goetz et al., 2010), it was expected that this reappraisal may increase positive affect as well. This was contrasted with a distancing strategy like that utilized in many intrapersonal emotion regulation studies, including the “collapse of compassion” studies (Cameron & Payne, 2011), in which participants attempted to decrease negative empathic affect, but reframe the situation in an avoidant way. It was hypothesized that while both reappraisal strategies may decrease negative affect, the hopeful reappraisal would be more successful than the distancing reappraisal in promoting costly helping behavior toward strangers, given its closer association with approach motivation and empathic concern, rather than avoidance of a source of personal distress. Based on research cited above, it was expected that moderate levels of negative affect may be most likely to promote prosocial behavior, in which there is a quadratic association between the two. Further, individual differences in reappraisal tendency may moderate the effect of instructed
reappraisal on prosocial behavior, such that those with a greater tendency to utilize reappraisal will show more prosocial behavior, particularly in a hopeful appraisal.

Additionally, the current set of studies sought to examine how reappraisal processes may contribute to prosociality in highly altruistic individuals. Altruistic kidney donors are known to have a heightened sensitivity to distress in others, as demonstrated by heightened amygdala activation in response to distress cues in the form of fearful faces in a passive viewing task (Marsh et al., 2014), and increased amygdala activation and functional connectivity with the periaqueductal gray in response to combined cues of distress and vulnerability described in Study 1. As introduced in Study 1, this highlights a potential paradox, in which highly altruistic individuals are particularly sensitive to distress cues across several paradigms, and demonstrate heightened neural responses to the pain of a stranger as described in Study 2, yet they have volunteered to undergo a major medical procedure to donate one of their own internal organs to a total stranger, with all of the associated risks and inconveniences. How such altruists regulate their emotional response to individuals in distress may explain how they can at once be extraordinarily in tune with distress yet engage in risky helping behavior that many would find to be distressing in itself. It was hypothesized that altruists would be more effective in implementing reappraisal to regulate empathic negative affect and that this distinction may be particularly pronounced for a hopeful reappraisal congruent with empathic concern.

The goals of the current set of studies were thus two-fold, explored in a single paradigm. First, the validity of an empathic emotion regulation task was tested in a population of healthy adults from the Georgetown community. In this task, participants viewed several sets of photos of distressed individuals in need, engaged in both reappraisal strategies in a counterbalanced order, and decided whether or not they would make costly donations to help the subjects of the
photos. Second, a sample of healthy adults from the larger Washington, D.C., community was compared to altruistic kidney donors as a sample of individuals known to be particularly sensitive to vulnerability and distress in others but who have also, seemingly paradoxically, engaged in a costly and risky helping behavior toward a stranger. In addition to the bias toward protective rather than preservative helping behaviors supported by an evolutionarily conserved mammalian offspring care system highlighted in Study 1, a greater tendency toward well-regulated empathic distress may promote the extraordinary helping behavior observed in this population.

**Study 3a**

**Methods**

**Participants.** Fifty-one healthy adults (34 female) took part in the study for monetary payment ($10) or course credit. The primary exclusion criterion was a lack of fluency in English, which resulted in the exclusion of one participant following consent, since the participant was unable to understand the task instructions. This resulted in a final sample of 50 participants (34 female) between 18 and 58 years old ($M = 21.62, SD = 6.01$). Of these 50 participants, 67% were Caucasian, 18% were Asian, 6% were Black or African-American, and 10% endorsed a mixed or not otherwise specified race. Most participants were currently enrolled undergraduates, while 18% had completed college and 10% had a graduate degree. All study procedures were approved by the Institutional Review Board at Georgetown University and all participants provided written informed consent before testing.

**Validation of negative photographs.** Neutral photos for the regulation task were selected from the previously validated Nencki Affective Picture System (NAPS; Marchewka, Zurawski, Jednorog, & Grabowska, 2013) photo database. Negative photos were selected from
NAPS and publicly available web sources, including major news organizations and charities. The negative photos selected all depicted a single individual in a context indicating a need for help (e.g., natural disaster, homelessness), many of whom were visibly expressing distress (e.g., crying).

Negative and neutral photos were validated in a separate participant sample via Amazon’s Mechanical Turk (mTurk). One hundred and nine healthy adults (68 female) from across the United States between 19 and 65 years old ($M = 35.43$, $SD = 12.27$) completed the validation task on mTurk. To maintain anonymity, no identifiable information was collected from participants. Consent was given by clicking to continue the task following a consent statement. Following brief instructions and practice with the response options, participants were required to correctly answer two short comprehension questions about how to respond to affect rating scales after each photo. To keep the duration of the validation task short, each participant rated a randomly selected 55 photo subset of a total set of 110 photos, which included 28 neutral and 82 negative photos, such that each photo received about 55 ratings (range: 53-58 ratings). Participants viewed each of the 55 photos for 6 seconds, rating both their positive and negative affect in response to each photo immediately following each 6 second presentation, on separate unipolar scales ranging from 0 (no pleasant/unpleasant feelings) to 8 (strong pleasant/unpleasant feelings). The use of two unipolar scales rather than one bipolar scale in self-reported affect has been shown to be more useful for allowing participants to express mixed affective responses (e.g., concurrent positive and negative feelings would result in a rating of 0 on a bipolar scale, rather than moderate to high ratings on each of two unipolar scales), and also to be a useful measure of affective arousal (whereas valence and arousal become dissociated for mixed feelings on a bipolar scale) (Kron, Goldstein, Lee, Gardhouse, & Anderson, 2013). Each scale was
presented as a slider, but only whole number responses were recorded. Keeping with Kron et al. (Kron et al., 2013), participants were instructed that a maximal rating on the pleasant scale represents feeling completely pleased, happy, satisfied, content, or hopeful; a minimal rating on the pleasant scale represents a neutral state of no pleasant feelings; a maximal rating on the unpleasant scale represents feeling completely unpleasant, unhappy, annoyed, unsatisfied, melancholic, or despaired; and a minimal rating on the unpleasant scale represents a neutral state of no unpleasant feelings. Affect rating scales were always presented in the same order (e.g., pleasant, then unpleasant) after each photo within each participant, counterbalanced in order across participants. Participants had unlimited time to respond to each affect scale. Data from one participant were dropped from analyses since the participant never moved the response slider (recording 0 for each affect scale for each photo).

**Empathic Emotion Regulation Task.** For the laboratory task, following consent, participants completed the empathic emotion regulation paradigm, in which they viewed a total of 10 neutral and 30 negative images and followed cognitive reappraisal prompts to modulate their affective response to each. Prior to beginning the empathic emotion regulation task, participants were given a brief overview of the experiment, and given a $10 endowment for the donation component of the task. The endowment was provided in the form of ten $1 bills that participants were instructed to put in their pocket or in their wallet, etc., so that it was within their possession. For the task, 36 negative photos and 10 neutral photos were selected based on the mTurk validation ratings, such that there were 10 photos per block to correspond to dollar amounts $1-10, and so that there were 3 practice photos per reappraisal condition. The negative photos were those rated as most negative and least positive, while neutral photos were selected as low in both negative and positive ratings, based on rankings of photos on average positive and
negative affect from the mTurk ratings (Table 3.1). The task was presented on an iMac desktop computer via Superlab version 4.0. All photos were presented in color, at the center of the screen, with a black background.

Table 3.1

<table>
<thead>
<tr>
<th>Photo Type</th>
<th>Pleasant Rating $M$ (SD)</th>
<th>Unpleasant Rating $M$ (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>2.75 (0.59)</td>
<td>1.51 (0.42)</td>
</tr>
<tr>
<td>Negative</td>
<td>0.61 (0.20)</td>
<td>5.78 (0.37)</td>
</tr>
</tbody>
</table>

The first part of the task was a practice block, in which participants practiced the reappraisal conditions (3 photos per practice block, for a total of 6 practice photos), responding to the affect scales, and responding to the donation requests. An experimenter checked with the participant following these practice trials and confirmed comprehension before proceeding with the task.

Participants viewed 30 unique negative photos, in blocks of 10 photos distributed across the three negative conditions. All three negative blocks were matched in positive and negative valence, based on the task validation data. Which negative photo block was associated with each condition was counterbalanced across participants, such that across all participants each block of 10 photos was presented equally often in each condition. Within each block, photos were presented in random order. In a within-subjects design, all participants completed blocks of view neutral, view negative, hope, and distance photo appraisals. The view neutral block was always first, followed by view negative. These two baseline blocks, in which participants viewed neutral or negative photos as they naturally would while maintaining their attention on them, were then followed by the two reappraisal blocks, in counterbalanced order across participants (Figure 3.1).
Figure 3.1. Task structure for the empathic emotion regulation task.

In each of the reappraisal blocks, participants viewed a set of negative photos, and attempted to down-regulate any experienced negative affect in either a hopeful or distancing way. In a hopeful appraisal, participants were instructed, “You may decrease negative feelings in a hopeful way by thinking things like, ‘It’s not a hopeless situation for this person,’ ‘Something could be done to make the situation better,’ or ‘I could help this person.’” In a distancing reappraisal, participants were instructed, “You may decrease negative feelings in a distancing way by thinking things like, ‘It is a hopeless situation for this person,’ ‘Nothing could be done to make the situation better,’ or ‘There’s nothing I can do.’” In this way, distancing was not simply a neutrally objective reappraisal in which participants simply viewed themselves as more distant from the subject of the photo, but rather was more actively avoidant of the subject in the photo, in order to contrast this strategy motivationally with the hope strategy. Such contradicting reappraisals were selected to interrogate potentially opposite effects of decreasing negative affect in response to another’s distress via cognitive reappraisal, depending on the specific appraisal, either eliciting an approach-oriented or avoidant-oriented motivational state, respectively (Figure 3.2).
Figure 3.2. Block structure for empathic emotion regulation task.

Each photo block began with a 4 second instruction to view or reappraise the upcoming photos. Following this was the sequence of 10 photos (Figure 3.2A). Each photo appeared for 8 seconds and was followed by two affect rating scales, each of which appeared for 4 seconds. The rating scales required participants to rate their own experiences of positive and negative affect, respectively, on a unipolar scale (Kron et al., 2013) that ranged from 0 (no pleasant/unpleasant feelings) to 8 (strong pleasant/unpleasant feelings). Participants responded via keyboard press with the number corresponding to their response. Affect rating scales were always presented in the same order within each participant, but counterbalanced in order across participants.

Immediately after each negative photo block was a donation task pertaining to the photos in that block (Figure 3.2B). Following brief instructions, the 10 photos just viewed or
reappraised were viewed again in random order for 8 seconds each. Each photo was followed by a randomly selected proposed donation amount that varied from $1 to $10 across the photos (with each value appearing once per block). Participants responded either yes or no (by pressing 1 or 0 keys, respectively) to donate the proposed amount to a charity that could help the person in the photo. Specifically, participants were instructed prior to beginning the task that one trial would be randomly selected at the end of the study and the money would go to the Red Cross, so that they should respond to each donation trial independently (i.e., they had $10 to “spend” on each trial). This donation paradigm is consistent with other recent investigations of the effects of affect on donation outcomes in response to vulnerable victims (Genevsky, Vastfjall, Slovic, & Knutson, 2013). A 30 second fixation separated each photo block (Figure 3.2C).

**Questionnaires.** Following the empathic emotion regulation task, participants completed self-report questionnaires. These questionnaires included state mood via the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), trait emotional reactivity via the Emotional Reactivity Scale (ERS; Nock, Wedig, Holmberg, & Hooley, 2008), trait emotion regulation via the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003), trait empathy via the Interpersonal Reactivity Index (IRI; Davis, 1980), basic demographic information, and self-reported psychological status including any psychiatric diagnoses, psychological symptoms, and current medications or drug use. The Emotion Regulation Questionnaire has both trait reappraisal and trait suppression subscales, reflecting the tendency to cognitively reframe emotional situations and the tendency to inhibit emotional expression, respectively, to modulate emotional responses. Given that the empathic emotion regulation task utilizes a cognitive reappraisal framework, only the reappraisal subscale was of interest. Similarly, the Emotional Reactivity Scale measures several aspects of emotional reactivity.
These include sensitivity, arousal-intensity, and persistence, which measure the ease with which an emotional reaction is elicited, the amplitude of the strength of an emotional reaction, and the duration of an emotional reaction, respectively. The Interpersonal Reactivity Index measures affective empathy via personal distress and empathic concern subscales, which are consistent with how these constructs have been discussed thus far, and cognitive empathy via perspective-taking and fantasy, with perspective-taking most closely measuring theory of mind. As a manipulation check and also to gather information on reappraisal execution, participants were also queried on the strategies they used in each part of the task, how difficult they found each emotion regulation strategy to be, and how successful they felt in utilizing each strategy.

**Analysis.** Donation choices were the main outcome of interest. Only the first valid response during donation choice events (the presentation of a dollar amount following a photo) was counted, resulting in an average donation response rate across the three blocks of 96.47% (range: 56.67-100.00%). One participant had no valid donation responses during the *view negative* block, but was retained in analyses not dependent on donation responses in this block (minimum participant donation response rate across conditions excluding this participant: 80%). Responses were analyzed through the generalized estimating equations (GEE) method of logistic regression in SPSS 24. GEE is a semiparametric analysis method that uses generalized linear models while accounting for correlated repeated measurements, thus allowing multiple responses within each condition for each participant. Response to each donation opportunity was the binomial response variable. Condition was a within-subjects predictor variable with three levels (*view negative, distance, hope*). Given the relatively large age range of participants, models were also tested with age as a covariate of no interest. An exchangeable working correlation matrix was specified, as correlations between repeated trials were expected to be equivalent. A model-
based estimator was used for the covariance matrix, since a subject variable was also specified, thus accounting for the repeated nature of within-subject measurements, and also given the relatively small sample size.

As with donation decisions, only the first valid response to each affect rating scale was counted, resulting in an average affect response rate of 94.60% (range: 77.50-100.00%). Mean affect ratings were calculated for each condition, separately for positive (pleasantness) and negative (unpleasantness) affect. Pleasantness and unpleasantness ratings were also summed as a measure of emotional arousal, per Kron and colleagues (Kron et al., 2013). Affect ratings across conditions were compared via repeated measures ANOVAs and paired samples t tests. Affect ratings for each photo individually were also considered in predicting trial-level donation behavior.

Bivariate correlations were examined between variables, with a focus on donation rates, self-reported affect, and trait reappraisal. Based on the reviewed hypotheses regarding potential nonlinear associations between affect and prosociality, quadratic associations were examined in addition to linear associations via hierarchical linear regression models with the addition of a quadratic term calculated by squaring the predictor of interest.

**Results**

Donation responses were a primary outcome of interest. Responses were analyzed through logistic regression in which the decision to donate on each trial was the binary response variable. First considering an effect of the three conditions preceding donation blocks (view negative, distance, hope) on donation behavior, there was a trend toward a main effect of condition, $\chi^2(2) = 5.59, p = .061$. Post hoc comparisons revealed that hope resulted in higher donation rates than distance, $p = .021$, though neither condition differed significantly from view
negative, both $p > .05$ (Figure 3.3). Condition remains a trend predictor of donation rates with age included as a covariate, $\chi^2(2) = 5.59, p = .061$. Adding donation amount as a continuous predictor and its interaction with condition to the model, there was a main effect of donation amounts, $\chi^2(1) = 92.56, p < .001$, in which donation rates decreased with increasing donation amount, but no interaction with condition, $\chi^2(2) = 2.28, p = .320$.

Figure 3.3. Donation rates by condition in Study 3a. Error bars represent 95% confidence intervals based on the SEM.

Given an expected moderating effect of trait reappraisal between condition and donation behavior, trait reappraisal on the ERQ and its interaction with condition were examined next. In a model with trait reappraisal tendency, condition, and their interaction predicting donation responses, a significant interaction between trait reappraisal and condition was observed, $\chi^2(2) = 8.50, p = .014$. This interaction remains when controlling for age, $\chi^2(2) = 8.42, p = .015$, and was specific to reappraisal tendency; no such interaction was observed for trait suppression tendency, $\chi^2(2) = 1.00, p = .606$. Examining correlations between trait reappraisal and donation rates in each condition confirms that trait reappraisal positively predicts donation rate in the hope
condition, \(r(48) = .28, p = .046\), but not the distance, \(r(48) = .09, p = .534\), or view negative, \(r(47) = .06, p = .692\), conditions (Figure 3.4).

![Figure 3.4. Correlations between trait reappraisal and donation rates by condition in Study 3a. Lines of best fit and 95% confidence intervals are displayed.](image)

Associations between self-reported affect and donation behavior were examined next. Examining negative affect and its interaction with condition predicting donation behavior, there was a main effect of negative affect, \(\chi^2(1) = 71.92, p < .001\), and a non-significant trend toward an interaction with condition, \(\chi^2(2) = 4.85, p = .089\). In a separate model with positive affect and its interaction with condition predicting donation behavior, there was both a main effect of positive affect, \(\chi^2(1) = 3.97, p = .046\), and a significant positive affect x condition interaction, \(\chi^2(2) = 10.33, p = .006\). In a third parallel model with arousal and its interaction with condition predicting donation behavior, there was a main effect of arousal, \(\chi^2(1) = 49.27, p < .001\), but no arousal x condition interaction, \(\chi^2(2) = 2.77, p = .251\). Finally, in a full factorial model with
positive affect, negative affect, condition, and their interactions predicting donation behavior, in order to test for a potential three way interaction, no such three way interaction was observed, $\chi^2(2) = 0.04, p = .980$.

Examining zero order correlations to interrogate the affect-related main effects and interactions described above, it was observed that while negative affect and arousal positively predict donation rates in each reappraisal condition, positive affect in isolation does not significantly predict donations in any condition (Table 3.2). No quadratic associations between positive or negative affect and donation rates were observed. Associations between self-reported affect and donation rates are displayed in Figure 3.5.

Table 3.2

*Bivariate Correlations between Donation Rates and Self-Reported Affect in Study 3a*

<table>
<thead>
<tr>
<th>Donation Rate</th>
<th>Pleasant Affect</th>
<th>Unpleasant Affect</th>
<th>Arousal Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hope</td>
<td>.11</td>
<td>.36**</td>
<td>.38**</td>
</tr>
<tr>
<td>Distance</td>
<td>-.23</td>
<td>.45***</td>
<td>.28*</td>
</tr>
<tr>
<td>View Negative</td>
<td>.05</td>
<td>.08</td>
<td>.10</td>
</tr>
</tbody>
</table>

*Note. *p < .05. **p < .01. ***p < .005.*
Figure 3.5. Correlations between positive and negative affect and donation rates by condition. Lines of best fit and 95% confidence intervals are displayed.

Affect ratings for photos in each condition were analyzed to examine the elicitation of negative affect and the effects of cognitive reappraisal on both negative and positive affect. A 2 (emotion) x 4 (condition) ANOVA revealed a main effect of condition, $F(3,47) = 15.13, p < .001, \eta^2_p = .491$, main effect of emotion, $F(1,49) = 203.68, p < .001, \eta^2_p = .806$, and condition x emotion interaction, $F(3,47) = 64.67, p < .001, \eta^2_p = .805$ (Figure 3.6). View neutral affect ratings were significantly more positive and less negative than all negative blocks, all $p < .001$, though the mean difference was smaller for positive affect ratings for hope and view neutral, $t(49) = 2.02, p = .049$. 
Next focusing on affect ratings in the three negative blocks, both hope, $t(49) = 4.18, p < .001$, and distance, $t(49) = 3.91, p < .001$, were rated as less negative than view negative, suggesting that cognitive reappraisal was successful in both conditions. Hope and distance did not differ significantly in their negative affect rating, $t(49) = 0.27, p = .790$, suggesting that with regard to negative affect the two reappraisals were equivalent. However, the two reappraisal conditions did differ with regard to positive affect ratings. While distance did not differ significantly from view negative in positive affect ratings, $t(49) = 0.94, p = .352$, hope was rated as significantly more positive than both view negative, $t(49) = 4.46, p < .001$, and distance, $t(49) = 4.29, p < .001$. See Figure 3.6 for mean positive and negative affect ratings per condition.

Next, the calculated affective arousal index was examined across conditions, supported by the main effect of condition on affect ratings as described above. As expected, view neutral had a lower arousal index than all negative blocks, all $p < .001$. While distance had a
significantly lower arousal index than both view negative, \( t(49) = 2.83, p = .007 \), and hope, \( t(49) = 2.95, p = .005 \), arousal for hope did not differ significantly from view negative, \( t(49) = 0.23, p = .821 \), as a result of the higher positive affect ratings for this reappraisal condition. See Figure 3.7 for mean arousal index per condition.

**Figure 3.7.** Index of affective arousal across conditions in Study 3a. Arousal was calculated by summing positive and negative affect intensity ratings, per Kron and colleagues (Kron et al., 2013). Error bars represent 95% confidence intervals based on the SEM.

**Summary.** In this first study of the empathic emotion regulation task, instructed reappraisal affected both self-reported affect and donation outcomes. Both reappraisal strategies decreased self-reported negative affect, but the hope strategy also increased positive affect. Despite equivalent decreases in negative affect, the two reappraisal conditions diverged in donation rates, with higher rates in the hope than distance condition, perhaps due to heightened positive affect. Trait reappraisal predicted donation rates in the hope condition specifically, suggesting that those with the greatest tendency to reappraise were most successful in using this reappraisal strategy to increase prosocial behavior. Given this moderating effect of trait reappraisal, if highly altruistic individuals are more proficient in regulating empathic distress,
then it is expected that this group may show an even larger effect of an approach-oriented appraisal on donation behavior in Study 3b.

**Study 3b**

**Methods**

**Participants.** Fifty-eight healthy adults between 21 and 60 years old \((M = 40.74, SD = 9.21)\) took part in the study for monetary payment (see Table 3.3 for participant characteristics). This sample included 30 altruistic kidney donors and 28 controls matched on most major demographic variables. Recruitment procedures were the same as those reported in Studies 1 and 2. Samples of altruists and controls were largely overlapping for Studies 2 and 3b. Twenty-nine altruists in this study also completed the empathic pain task in Study 2; one altruist in this study was excluded from the neuroimaging component due to an age over 55 years. Twenty-six controls in this study also completed the empathic pain task in Study 2 (two controls from Study 2 did not complete Study 3b due to equipment error); the remaining 2 controls that completed Study 3b but not Study 2 participated following the conclusion of the neuroimaging component of the study battery. Since this study was in the same task battery as Study 2, exclusion and inclusion criteria were the same, with one exception for a participant over 55 years of age. All study procedures were approved by the Institutional Review Board at Georgetown University and all participants provided written informed consent before testing.
Table 3.3

Study 3b Participant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Altruists (n = 30)</th>
<th>Controls (n = 28)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female (% Male)</td>
<td>11/19 (36.7%)</td>
<td>12/16 (42.9%)</td>
<td>.630</td>
</tr>
<tr>
<td>White/Other Race (% White)</td>
<td>27/3 (90.0%)</td>
<td>25/3 (89.3%)</td>
<td>.929</td>
</tr>
<tr>
<td>Household Income ≥ $60,000</td>
<td>20 (71.4%)</td>
<td>25 (92.6%)</td>
<td>.056</td>
</tr>
<tr>
<td>Education ≥ Four-Year Degree</td>
<td>20 (66.7%)</td>
<td>26 (92.9%)</td>
<td>.014</td>
</tr>
<tr>
<td>Age M (SD)</td>
<td>42.70 (10.03)</td>
<td>38.64 (7.89)</td>
<td>.094</td>
</tr>
<tr>
<td>IQ M (SD)</td>
<td>107.45 (12.63)</td>
<td>111.96 (11.39)</td>
<td>.162</td>
</tr>
</tbody>
</table>

Note. One control did not report their household income. One altruist declined the KBIT.

Procedures. The empathic emotion regulation task and questionnaires in Study 3b were identical to Study 3a. The primary procedural differences from Study 3a were the recruitment of altruistic kidney donors and matched controls in Study 3b, and the administration of the task and questionnaires in the context of a larger battery of behavioral tasks and questionnaires following fMRI scanning in Study 3b. All participants completed the same battery of behavioral tasks and questionnaires. Twenty-nine altruists and 26 controls in this study also completed the neuroimaging component, which was partially reported in Study 2.

Analysis. Donation choices were again the main outcome of interest. Only the first valid response during donation choice events (the presentation of a dollar amount following a photo) was counted, resulting in an average donation response rate across the three blocks of 97.93% (range: 83.33-100.00%). Given group differences in education level, models were also tested with education level as covariate of no interest. As in Study 3a, responses were analyzed through GEE logistic regression. Response to each donation opportunity was the binomial response variable. Condition was a within-subjects predictor variable with three levels (view negative, distance, hope) and group was a between-subjects variable. An exchangeable working correlation matrix and model-based estimator were again specified.
As with donation decisions, only the first valid response to each affect rating scale was counted, resulting in an average affect response rate of 94.14% (range: 80.00-100.00%). Mean affect ratings were calculated for each condition, separately for positive and negative affect. Pleasantness and unpleasantness ratings were also summed as a measure of emotional arousal, per Kron and colleagues (Kron et al., 2013). Affect ratings across conditions were compared via repeated measures ANOVAs and paired samples t tests. Affect ratings for each photo individually were also considered in predicting trial-level donation behavior.

The goals for analyses for Study 3b were to examine any group differences in the effects of condition, self-reported affect, and trait reappraisal on donation behavior. The analysis approach followed that of Study 3a, with the addition of tests for group differences. Affect ratings across conditions were compared via repeated measures ANOVAs and paired samples t tests. Comparisons between groups were conducted via mixed model ANOVAs and independent samples t tests. Bivariate correlations were examined between variables, with a focus on donation rates, self-reported affect, and trait reappraisal. Based on the reviewed hypotheses regarding potential nonlinear associations between affect and prosociality, quadratic associations were examined in addition to linear associations via hierarchical linear regression models with the addition of a quadratic term calculated by squaring the predictor of interest.

**Results**

Donation responses were again a primary outcome of interest. Responses were analyzed through logistic regression in which the decision to donate on each trial was the binary response variable. First, a model was tested with appraisal condition, group, and their interaction predicting donation behavior. While there was a main effect of group, $\chi^2(1) = 4.53, p = .033$, in which altruists had higher donation rates overall, $p = .028$, there was no main effect of condition,
\( \chi^2(2) = 1.22, p = .543 \), nor a group x condition interaction, \( \chi^2(2) = 4.04, p = .133 \) (Figure 3.8). This pattern of results remained largely the same while controlling for education level, though the main effect of group was reduced to trend-level, \( \chi^2(1) = 3.01, p = .083 \). Adding donation amount as a continuous predictor and its interactions with condition and group to this model, there is again a main effect of donation amount, \( \chi^2(1) = 75.67, p < .001 \), in which higher donation amounts predict lower donation rates, but no interactions with condition or group, all \( p > .05 \). This pattern of results is the same if education level is added as a covariate.

![Figure 3.8. Donation rates by group and condition in Study 3b. Error bars represent 95% confidence intervals based on the SEM.](image)

With trait reappraisal and its interactions with group and condition considered next, it was found that while there was a main effect of trait reappraisal on donation rates, \( \chi^2(1) = 8.24, p = .004 \), there was no interaction with condition, \( \chi^2(2) = 2.24, p = .327 \), group, \( \chi^2(1) = 0.18, p = .669 \), nor with condition x group, \( \chi^2(2) = 0.56, p = .756 \). This pattern of results was the same when education level was added as a covariate.
Associations between self-reported affect and donation behavior were examined next. Examining negative affect and its interactions with group and condition predicting donation behavior, there was a main effect of negative affect, $\chi^2(1) = 80.03, p < .001$, and an interaction with group, $\chi^2(1) = 6.70, p = .010$, but no other interactions. In a separate model with positive affect and its interactions with group and condition predicting donation behavior, there was a main effect of positive affect, $\chi^2(1) = 6.44, p = .011$, and a non-significant trend toward an interaction between positive affect and group, $\chi^2(1) = 2.88, p = .090$, but no other interactions. In a third parallel model with arousal and its interactions with group and condition predicting donation behavior, there was a main effect of arousal, $\chi^2(1) = 38.51, p < .001$, but no interactions. Finally, in a full factorial model with positive affect, negative affect, group, condition, and their interactions predicting donation behavior, in order to test for a potential four way interaction, no such interaction was observed, $\chi^2(2) = 1.35, p = .509$.

Examining zero order correlations to interrogate the affect-related main effects and interactions described above, similar patterns were observed in each group. While negative affect and arousal positively predict donation rates in each condition, positive affect does not significantly predict donations in any condition (Table 3.4). No quadratic associations between positive or negative affect and donation rates were observed. Associations between self-reported affect and donation rates are displayed in Figure 3.9.
Table 3.4

*Bivariate Correlations between Donation Rates and Self-Reported Affect in Study 3b*

<table>
<thead>
<tr>
<th></th>
<th>Donation Rate</th>
<th>Pleasant Affect</th>
<th>Unpleasant Affect</th>
<th>Arousal Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hope</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altruists</td>
<td>.12</td>
<td>.41*</td>
<td>.32</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>-.09</td>
<td>.61****</td>
<td>.50**</td>
<td></td>
</tr>
<tr>
<td><strong>Distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altruists</td>
<td>.13</td>
<td>.73****</td>
<td>.59****</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>-.23</td>
<td>.73****</td>
<td>.60****</td>
<td></td>
</tr>
<tr>
<td><strong>View Negative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altruists</td>
<td>.04</td>
<td>.66****</td>
<td>.53***</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>-.35</td>
<td>.77****</td>
<td>.66****</td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05. **p < .01. ***p < .005. ****p < .001.
Figure 3.9. Correlations between affect and donation rates in each condition, by group. Lines of best fit and 95% confidence intervals are displayed.
Groups similar in trait measures and state mood. As summarized in Table 3.5, altruists and controls did not differ significantly in affective trait measures, or in state positive and negative affect as measured by the PANAS, though there was a trend toward higher empathic concern for altruists, $p = .059$. Altruists and controls were also similar on their perceptions of success and difficulty in applying the reappraisal strategies, although altruists rated the distance strategy as more difficult than controls, $t(56) = 2.35, p = .022$.

Table 3.5

<table>
<thead>
<tr>
<th>Study 3b Trait Measures and State Mood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Altruists $(n=30)$</th>
<th>Controls $(n=28)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERQ Reappraisal</td>
<td>30.80 (6.30)</td>
<td>32.14 (6.17)</td>
<td>.416</td>
</tr>
<tr>
<td>ERQ Suppression</td>
<td>13.13 (4.26)</td>
<td>12.50 (4.99)</td>
<td>.604</td>
</tr>
<tr>
<td>ERS Sensitivity</td>
<td>11.00 (7.22)</td>
<td>8.18 (5.11)</td>
<td>.093</td>
</tr>
<tr>
<td>ERS Arousal-Intensity</td>
<td>10.07 (5.28)</td>
<td>8.14 (4.34)</td>
<td>.137</td>
</tr>
<tr>
<td>ERS Persistence</td>
<td>5.43 (3.49)</td>
<td>4.82 (2.61)</td>
<td>.455</td>
</tr>
<tr>
<td>IRI Empathic Concern</td>
<td>22.60 (4.64)</td>
<td>20.21 (4.78)</td>
<td>.059</td>
</tr>
<tr>
<td>IRI Personal Distress</td>
<td>8.90 (4.71)</td>
<td>10.00 (6.12)</td>
<td>.445</td>
</tr>
<tr>
<td>IRI Perspective-Taking</td>
<td>20.77 (4.05)</td>
<td>19.29 (4.18)</td>
<td>.176</td>
</tr>
<tr>
<td>IRI Fantasy</td>
<td>16.77 (6.62)</td>
<td>15.43 (7.43)</td>
<td>.471</td>
</tr>
<tr>
<td>PANAS Positive Affect</td>
<td>32.40 (7.18)</td>
<td>31.11 (7.19)</td>
<td>.496</td>
</tr>
<tr>
<td>PANAS Negative Affect</td>
<td>11.83 (3.42)</td>
<td>11.32 (2.06)</td>
<td>.497</td>
</tr>
</tbody>
</table>

Note. Means and standard deviations are reported. ERQ = Emotion Regulation Questionnaire, ERS = Emotional Reactivity Scale, IRI = Interpersonal Reactivity Index, PANAS = Positive and Negative Affect Schedule.

Reappraisal decreases negative affect. There was no evidence of a 2 (emotion) x 4 (condition) x 2 (group) interaction in affect ratings, $F(3,54) = 0.19, p = .903, \eta^2 = .010$.

However, main effects of condition, $F(3,54) = 30.01, p < .001, \eta^2 = .625$, emotion, $F(1,56) = 150.39, p < .001, \eta^2 = .729$, and a condition x emotion interaction, $F(3,54) = 154.06, p < .001, \eta^2 = .895$, were observed. View neutral affect ratings were significantly more positive and less negative than all negative blocks, all $p < .001$.  

88
Next, focusing on affect ratings in the three negative blocks, both hope, $t(57) = 5.31, p < .001$, and distance, $t(57) = 5.41, p < .001$, were rated less negative than view negative, suggesting that cognitive reappraisal was successful in both conditions. Hope and distance did not differ significantly in their negative affect rating, $t(57) = 0.67, p = .506$, suggesting that with regard to negative affect the two reappraisals were equivalent. Both distance, $t(57) = 2.91, p = .005$, and hope, $t(57) = 5.92, p < .001$, were rated as significantly more positive than view negative.

However, hope was rated as more positive than distance, $t(57) = 5.17, p < .001$. Thus, across the Study 3b sample, effects of condition on affect were replicated, with an additional finding of increased positive affect in the distance condition, though no group differences emerged. See Figure 3.10 for affect ratings by group across conditions.

![Figure 3.10](image_url)

*Figure 3.10.* Affect ratings across conditions for altruists and controls. Error bars represent 95% confidence intervals based on the SEM.

Affective arousal was explored next. There was no evidence of a 4 (condition) x 2 (group) interaction in affective arousal, $F(3,54) = 1.41, p = .250, \eta^2_p = .073$. The main effect of
condition described above was examined through post hoc comparisons, confirming findings from Study 3a, in which view neutral was characterized by lower arousal than all negative conditions, all \( p < .05 \), view negative and distance did not differ, \( t(57) = 1.10, p = .278 \), but hope had higher arousal than both view negative, \( t(57) = 2.16, p = .035 \), and distance, \( t(57) = 4.61, p < .001 \) (Figure 3.11).

![Bar chart](image)

*Figure 3.11. Index of affective arousal across conditions for altruists and controls. Error bars represent 95% confidence intervals based on the SEM.*

**Summary.** While altruists were more prosocial overall, their donation behavior was not significantly affected by reappraisal instructions. Further, trait reappraisal did not interact with condition or group in predicting donation rates. However, group did interact with self-reported affect in predicting donation rates. Specifically, while positive affect tended to have negative associations with donation rates for controls, positive affect tended to have positive associations with donation rates for altruists. And, positive associations between negative affect and donation
rates tended to be stronger for controls than altruists. Groups did not otherwise differ on affective or empathic measures.

**Discussion**

The set of findings in this empathic emotion regulation task across two studies further supports a role for emotion regulatory appraisals in shaping empathic emotion and consequent costly prosociality, while also suggesting roles for both negative and positive affect in donation behavior. In Study 3a, combined increases in donation behavior during a hopeful appraisal and decreases in donation behavior during a distancing appraisal revealed a divergence of these two strategies with regard to prosocial outcomes. Further, there was a moderating effect of trait reappraisal on donation behavior, in which reappraisal tendency predicted prosociality in the *hope* reappraisal condition in particular. In Study 3b, altruists were more prosocial overall, though no effects of reappraisal condition emerged. Instead, interactions between group and condition suggest a role for a mix of negative and positive affect in predicting prosociality in altruists. While quadratic associations between empathic negative affect and prosociality were not supported by the current studies, together the findings reported here suggest that regulated empathic affect still has a role to play in prosociality, though cognitive reappraisal as traditionally defined may not be an explanatory variable in extraordinary altruism.

Contrary to hypotheses, quadratic effects of negative affect and arousal in predicting in prosociality were not observed. Rather, linear effects of positive and negative affect were found. The finding that negative empathic affect positively predicted donation behavior within each condition in almost all cases is consistent with heightened affective resonance with distress in extraordinary altruists observed in Studies 1 and 2, and also with theories of altruism emphasizing the role for concern in promoting prosociality (e.g., Nichols, 2001; Batson, 1991).
Though participants were not instructed to increase positive emotions, this was a byproduct of hopeful reappraisal. It may be that this reappraisal in particular promoted prosociality by a simultaneous decrease in negative affect and increase in positive affect. Such an effect is consistent with recent findings in the compassion literature, in which compassion training increases positive affect in response to distress in others and also promotes prosocial behavior (Goetz et al., 2010; Singer & Klimecki, 2014). Thus, it may be that a hopeful reappraisal resulted in a compassionate orientation. Such findings are consistent with a role for positive emotions in low cost helping and cooperation (Genevsky, Vastfjall, Slovic, & Knutson, 2013; Hauser, Preston, & Stansfield, 2014; Rand, Kraft-Todd, & Gruber, 2015). While there is a risk that repeated affect ratings after each photo could have resulted in less accurate ratings (i.e., if participants were satisficing with their responses) or that such ratings could have affected emotional responses, given the known effects of affect labeling on behavioral and neural affective responses (Lieberman et al., 2007), this method was selected as preferable to gathering affect ratings following the task such that affect would be measured upon first exposure to each photo and in the context of each appraisal block. In this way, affect ratings for each individual photo during each appraisal or reappraisal could be utilized to predict donation behavior in response to that specific photo in the applied analytical models.

In particular, this pattern of increased positive affect as a result of a hopeful appraisal across samples suggests connections to recent research exploring compassion as an emotion regulatory process supporting empathy and prosociality. Across several studies, Klimecki and colleagues (2013; 2014) have found that while baseline responses to distress in others tend to be characterized by negative affect and neural activation in regions associated with empathy for pain, particularly the anterior insula and anterior medial cingulate, compassion training increases
positive affect and neural activation in regions associated with affiliation and positive emotional experiences, including the medial orbitofrontal cortex and ventral striatum, while leaving negative affect intact. Further, empathy training focused on resonating with others’ suffering increased negative affect and associated neural activation. The authors (Klimecki et al., 2014) posit that this later finding suggests an antagonism between empathy (here, closer in meaning to personal distress than empathic concern) and compassion as separable neural systems.

The findings of Klimecki and colleagues (2013; 2014) inspired them to label compassion a “new coping strategy” in the face of potential personal distress in response to a needy other. Engen and Singer (2015) followed up on this interpretation and directly compared a compassionate orientation to traditional cognitive reappraisal, in a recent study of empathic responses to distress by experts in compassionate meditation. Defining compassion as maintaining a positive emotional state in response to suffering and reappraisal as imagining a more positive ending to the distressing situation, they found that instructed compassion primarily increases positive affect while instructed reappraisal primarily decreases negative affect, with expected diverging modulations of neural activation. Though they did not measure behavioral outcomes of these two appraisals, the authors hypothesized that while compassion would have the expected positive effects on prosociality, reappraisal as they defined it could lead to apathy.

Given that the reviewed studies of compassion were mostly neutral with regard to effects on negative affect, yet a hopeful reappraisal that promoted prosociality in the current study was characterized by concurrent increases in positive affect and decreases in negative affect, it remains to be seen what combined effects of regulatory appraisal processes on these independent emotional trajectories. Findings from the current study are consistent with the effects of mindfulness on empathic emotion, in which mindfulness simultaneously increased positive
emotion and decreased negative emotion during helping (Cameron & Fredrickson, 2015), further suggesting that decreasing personal distress and increasing empathic concern may be separable processes that can be influenced independently, rather than a single entity to be targeted. These studies, along with the current study, also highlight the methodological value of independently measuring positive and negative affect in order to test and manipulate these independent affective responses (Kron et al., 2013).

Several methodological limitations may have contributed to a lack of an effect of condition on donation behavior in Study 3b. Replicating results from Study 3a to Study 3b may have been limited by the task load from the larger study battery, especially since reappraisal is known to be an effortful strategy (Sheppes, Catran, & Meiran, 2009). Such differences may be why effects of reappraisal strategies on donation rates were absent in this second sample, though effect sizes were also small in the first study, highlighting a need for further investigation of the effects of emotion regulation on prosociality. The relatively low cost of prosociality in these studies is another potential limitation. While the donation task was designed to be consistent with prior studies of the effects of negative and positive affect on donation behavior in response to vulnerable victims (Genevsky, Vastfjall, Slovic, & Knutson, 2013) and also the experience of donating to a charity in the real-world, it is possible that participants did not view the donation stipend as their own or did not view the donation task as costly. However, there was variation in donation behavior, in which some participants never donated, and also donation behavior varied by condition, suggesting the nature of the task was not a significant methodological weakness. Yet, it is likely that stronger contrasts between altruistic kidney donors and non-donors may have been revealed if the stakes of helping were higher, given the very costly behavior that differentiates between the two groups.
Emotional processes not captured by trait measures of reappraisal tendency or instructed reappraisal in the context of low-cost prosociality may better characterize extraordinarily altruistic individuals. In particular, it may be that altruists tend to experience a mix of negative and positive affect in response to distress in others, rather than more simply regulated negative affect. While such an effect is more consistent with the compassion literature cited above and also the group by emotion interaction effects on donation behavior observed here, such patterns may not be captured by relatively coarse measures of trait emotion regulation tendency or explicit instructed reappraisal to down-regulate negative affect. Altruists and controls in Study 3b did not differ in self-reported affect in any study condition, however group by affect interactions in predicting prosociality suggest that affect differentiated between the groups in predicting donation behavior, regardless of condition.

In sum, cognitive reappraisal can promote low-cost prosociality, though the goals of such reappraisal are likely important for most individuals. Consistent with recent research on compassionate responding, while a hopeful appraisal may promote prosocial behavior, a distancing appraisal could lead to apathy and inaction. The current study adds to accumulating evidence that even short-term compassion training can increase non-reciprocal prosocial behavior toward strangers (Leiberg, Klimecki, & Singer, 2011) and that targeting both negative and positive aspects of the empathic response may be most effective in supporting affiliation with distressed others and consequent helping (Engen & Singer, 2015). Further research into both the independent and interacting effects of negative and positive empathic affect on prosociality will further clarify these patterns.
GENERAL DISCUSSION

While costly help extended to strangers is undoubtedly a complex psychological process with myriad affective and cognitive influences represented across multiple neural networks, the three studies presented in this dissertation represent steps toward understanding the empathic processes that drive care-based altruism. The goals of this dissertation were to answer three major questions regarding affective and cognitive processes that support empathy and altruism. First, does the amygdala-mediated distress sensitivity previously observed in extraordinary altruists contribute to caring decision-making? Study 1 found that amygdala-mediated distress sensitivity supports caring decision-making in extraordinary altruists, in coordination with the periaqueductal gray, suggesting a role for the mammalian offspring care system in human altruism. The role of this network in responding to cues of distress and vulnerability was supported by both functional and structural connectivity between these regions. Second, in addition to enhanced sensitivity to distress in others, do altruists also demonstrate enhanced similarity in first-hand and vicarious experiences of pain? Study 2 found that the role of heightened distress sensitivity in altruists may be further supported by enhanced affective resonance. Altruists demonstrated greater neural overlap for the first-hand experience of pain and the observation of pain in stranger, particularly in anterior insula. Further, this group-level resonance was supported by individual-level covariation in activation levels, and also greater matching in degree of activation between first-hand and vicarious pain in left anterior insula. Third, are emotion regulatory processes important for promoting well-regulated empathic concern and prosocial helping? Study 3 supported a role for empathic emotion regulation in supporting prosociality, but also highlighted that this may not be a strong explanatory variable in extraordinary altruists. While Study 3a suggests that instructed appraisals affect prosociality
through modulated affect, with a moderating role for trait reappraisal, Study 3b did not replicate these effects of condition and instead revealed that affect may have varying effects by group regardless of condition. Together, these studies support the roles of both affective and cognitive processes in creating a multidetermined concern, which will now be reviewed in more detail.

Studies 1 and 2 corroborate and extend past research findings that extraordinary altruists are characterized by enhanced neurocognitive sensitivity to distress in others. Building on the amygdala-mediated sensitivity to fearful facial expression in altruists reported previously (Marsh et al., 2014), amygdala connectivity with the periaqueductal gray was implicated in altruists’ caring decision-making in response to combined cues of distress and vulnerability in Study 1. These findings based in behavior, function, and structure implicate activity in the mammalian offspring care system in supporting compassionate responding and altruism (Preston, 2013; Simon-Thomas et al., 2012). Hrdy (2011) argues that the development of alloparenting, meaning the extension of care to vulnerable others beyond one’s own offspring, as a critical caregiving strategy led to the complexity of human empathy. Humans came to not only care for vulnerable others beside their own offspring, but also came to be particularly attuned to cues indicating the needs and intentions of others, which was an important skill for both caregivers and young seeking care from multiple sources. Such an offspring care system may be one mechanism that biases an individual toward protective other-regarding behavior rather than preservative self-interested behavior, characterized by the override of avoidance and withdrawal behaviors governed by the periaqueductal gray in response to amygdala-mediated responses to threat cues.

Study 2 represented a step forward in understanding what such amygdala-mediated distress sensitivity might suggest about affective resonance. While only the observation of cues of distress and vulnerability had been tested up to this point in altruists, Study 2 allowed for a
direct comparison between the observation and first-hand experience of painful pressure
stimulation through an empathic pain paradigm. Study 2 thus allowed for a more direct test of
simulation theories of empathy (Preston & De Waal, 2002), which would suggest that amygdala-
mediated distress sensitivity represents simulation of the fear experience in the self. This
empathic pain paradigm revealed that altruists do in fact demonstrate enhanced simulation of
pain, particularly in left anterior insula. Such a conclusion is supported by self-other covariation
in activation levels in this region of interest in altruists but not controls, which clarifies the
group-level neural overlap findings that could be limited in representing individual-level
correspondence in neural activation (Beckes et al., 2012).

Exactly what this self-other correspondence in anterior insula activation represents is still
an area of debate. While univariate analyses such as those reported here have suggested
meaningful self-other overlap in insula activation, which has been interpreted as a simulation of
the affective qualities of pain since somatosensory regions did not show the same sort of overlap
(Lamm et al., 2011), more recent multivoxel pattern analyses have produced mixed conclusions
regarding the function of the anterior insula in pain simulation. Specifically, Krishnan and
colleagues (2016) claim that neural signatures of first-hand and vicarious pain are distinct, while
Corradi-Dell’Acqua and colleagues (2016) have found at least some shared patterns in anterior
insula. Altruists are known to be sensitive to distress cues in particular (fear) and not negative
emotional cues in general (anger) (Marsh et al., 2014). Thus, comparing the data presented here
on self-other correspondence in pain processing in anterior insula to self-other correspondence in
fear processing during the anticipation period in this same paradigm could potentially speak to
the specificity of the anterior insula in distress simulation. If the insula shows self-other
correspondence in both conditions, and also if altruists show more overlap than controls in both
of these conditions, such findings would suggest that the role of the anterior insula may be more
domain-general for distress. The degree of anterior insula activation in each condition could also
be compared. It could be that the anterior insula responds to both instances of distress, but
responds more strongly to pain-related distress. The present analyses could also be extended
beyond region of interest approaches. Particularly given its role as an integrative center of
affective interoceptive information (Craig, 2003; Craig, 2009; Lamm & Singer, 2010), and
potential differences in interoceptive processing suggested by the group difference in pain
thresholds in Study 2, functional connectivity analyses with the anterior insula during the self
and other pain conditions could also reveal more information about the insula’s role in the
processes expected to be important for empathic pain, as part of a neural network rather than a
region in isolation. Specifically, it is hypothesized that the anterior insula and anterior cingulate
cortex form a hub with orbitofrontal cortex, ventral striatum, amygdala, and periaqueductal gray
that may be important for evaluating the motivational value of internal and external stimuli,
including social stimuli (Bernhardt & Singer, 2012).

Study 3 suggests that altruists may differ from controls in the interaction of bottom-up
and top-down empathic processes, but not as hypothesized at the outset of this dissertation.
Altruists and controls did not differ in affect levels or donation rates by reappraisal condition, but
affect did have different overall effects in predicting donation behavior by group. In particular,
while there were main effects of both negative and positive affect, negative affect interacted with
group such that it appeared to be less of a positive predictor for altruists and positive affect
trended toward an interaction with group, in that it appeared it may be a relatively positive
predictor for altruists but a relatively negative predictor for controls. Vaish argues that it is the
influence of top-down processes that makes empathic responding flexible (2016). It may be that
what sets altruists apart is that they rely less on this top-down influence, which could bias someone either toward or away from prosocial responding, so their responding is less flexible, in that they tend toward prosociality. Thus, it could be that altruists’ responding to others is not as dependent on situational factors, which in Study 3b would be the various conditions, resulting in higher prosociality across the board. Alternatively, it could be that altruists’ affective resonance is not dominated by negative affect, but is instead a mix of negative and positive affect, which would technically only be partially resonant with the distress of the target, but would be consistent with a compassion view of empathic concern rather than a regulated personal distress view of empathic concern. Klimecki and colleagues, in their studies of compassionate orientations on affiliation and prosocial responding (2014; 2015), suggest that there may be an antagonism between distress and care systems, consistent with prior views on personal distress and empathic concern. However, the association between these two systems is likely more complicated, given evidence including higher quality offspring care in rats high in anxiety relative to those low in anxiety (Bosch, 2011). While their studies support the influence of compassion through the elicitation of positive affect, they are relatively ambivalent about compassion’s effects on negative affect. Given the results summarized in Study 3, combined decreases in negative affect and increases in positive affect may be effective in promoting prosociality and altruism. While this recent research in the effects of compassionate orientations on empathy and prosocial behavior have begun to highlight the role of positive approach-related emotions, rather than just the down-regulation of negative affect, how such positive affect interacts with resonant negative affect to optimally promote prosociality is still unclear.
Conclusions

Together, the results of the current dissertation fit well within a process model of empathy (Decety et al., 2016), in which perceived distress leads to affective resonance which in turn promotes prosocial drive and potential helping behavior, with modulating factors including emotion regulation as a lever between affective resonance and prosocial drive. These findings also suggest that the interactions between relatively bottom-up affective resonance and relatively top-down emotion regulation in producing a “multidetermined concern” (Vaish, 2016) may differ between extraordinary altruists and controls. While the results of distress sensitivity and affective resonance suggest that these affective processes may be heightened in extraordinary altruists, the results of empathic emotion regulation suggest a potential role for reappraisal in promoting low-cost prosociality through the modulation of negative affect, though emotion regulation as measured here does not seem to explain prosociality in altruists. However, that study did highlight that altruists may differ from the average person in their affective reactions to distress in strangers, regardless of instructed appraisal. Recent findings in the compassion literature suggest potential avenues for future work to clarify the role of appraisals and interacting positive and negative affect in contributing to prosocial actions in such altruists.
APPENDIX

The care-eliciting scenarios task administered during fMRI scanning in Study 1 was previously validated in a separate sample of 55 healthy adults (39 female), age 18-25 ($M = 19.82$, $SD = 1.42$). Participants completed sequences of preattentive facial expression/lexical decision-making tasks and read and responded with sympathy ratings to brief scenarios. Each participant read 40 scenarios, each describing a stranger in a situation designed to be care-eliciting or not. Scenarios were adapted from an earlier behavioral version of this paradigm (Marsh & Ambady, 2007). Scenarios were followed by question blocks, in which participants reported how much sympathy they felt for the protagonist on a 1 to 7 scale. Prior to viewing each scenario, participants were presented with preattentive fearful and neutral facial expressions in the context of a lexical decision task (deciding whether a string of letters formed a real or nonsense word via button press—this task was included to provide a cover story for the emotional expressions). Word type was presented randomly. Affectively neutral letter strings were selected from the ANEW database of words (Bradley & Lang, 1999). Each string of letters was preceded by either a fearful or neutral facial expression that was presented for 25 milliseconds (Dimberg & Öhman, 1996) and immediately backwards masked by a 75 millisecond face scramble. Expressions were drawn from the Pictures of Facial Affect stimulus set (Ekman & Friesen, 1976). Participants completed blocks of eight facial expression sequences before each scenario. All expressions within each block were either fearful or neutral. Letter string, scenario, and question durations were determined by participant response latencies. This task structure resulted in two sets of 40 scenarios, half of which were care-eliciting and half of which were distractors, and each scenario type was further divided equally by preceding expression type. Each participant read 40 scenarios, with half of the participants completing each version, resulting in ratings for the 80
total scenarios that were presented to fMRI participants. Presentation of the four scenario types was randomly ordered.

There was a main effect of expression, $F(1, 54) = 35.66, p < .001$, $\eta^2 = .398$, main effect of scenario, $F(1, 54) = 256.39, p < .001$, $\eta^2 = .826$, and an expression x scenario interaction, $F(1, 54) = 114.65, p < .001$, $\eta^2 = .680$. Sympathy ratings were significantly higher for care-eliciting scenarios than for neutral scenarios, and sympathy ratings were also higher for scenarios preceded by fearful expressions than by neutral expressions (Table A.1). Importantly, sympathy ratings were higher for care-eliciting scenarios preceded by fearful expressions than care-eliciting scenarios preceded by neutral expressions, $t(54) = 11.46, p < .001$.

<table>
<thead>
<tr>
<th>Expression Type</th>
<th>Scenario Type</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Care-Eliciting</td>
<td>4.78</td>
<td>0.77</td>
</tr>
<tr>
<td>All</td>
<td>Distractor</td>
<td>2.99</td>
<td>0.79</td>
</tr>
<tr>
<td>Fear</td>
<td>All</td>
<td>4.09</td>
<td>0.71</td>
</tr>
<tr>
<td>Neutral</td>
<td>All</td>
<td>3.68</td>
<td>0.72</td>
</tr>
<tr>
<td>Fear</td>
<td>Care-Eliciting</td>
<td>5.29</td>
<td>0.81</td>
</tr>
<tr>
<td>Neutral</td>
<td>Care-Eliciting</td>
<td>4.27</td>
<td>0.87</td>
</tr>
<tr>
<td>Fear</td>
<td>Distractor</td>
<td>2.89</td>
<td>0.82</td>
</tr>
<tr>
<td>Neutral</td>
<td>Distractor</td>
<td>3.09</td>
<td>0.90</td>
</tr>
</tbody>
</table>

In this validation sample, three subjects reported a diagnosis of a mood or anxiety disorder (1 generalized anxiety disorder, 1 anxiety and depression, 1 depression), two of which were on psychiatric medications (paroxetine and fluoxetine), and a fourth subject reported current use of a narcotic painkiller (oxycodone). The pattern of results remains the same after the removal of these four subjects. This task was adapted for fMRI scanning, such that four expressions were presented prior to each scenario during the fMRI task, rather than eight.
REFERENCES


Progress in Neurobiology, 46(6), 575-605.


Cameron, C. D., & Fredrickson, B. L. (2015). Mindfulness facets predict helping behavior and


representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. Nature Communications, 7, 10904. doi:10.1038/ncomms10904


Decety, J., Bartal, I. B., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of
prosocial behaviour: Highly conserved neurobehavioural mechanisms across species.


doi:10.1002/9780470479216.corpsy0970


doi:10.1097/TLD.0000000000000040


Eisenberg, N., & Okun, M. A. (1996). The relations of dispositional regulation and emotionality to elders’ empathy-related responding and affect while volunteering. *Journal of*


Keltner, D., Kogan, A., Piff, P. K., & Saturn, S. R. (2014). The sociocultural appraisals, values,
and emotions (SAVE) framework of prosociality: Core processes from gene to meme.


*Neuropsychologia, 47*(10), 2073-2081. doi:10.1016/j.neuropsychologia.2009.03.017


*Journal of Neuroscience, 25*(49), 11489-11493. doi:10.1523/JNEUROSCI.3984-05.2005


Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and


119


Tovote, P., Esposito, M. S., Botta, P., Chaudun, F., Fadok, J. P., Markovic, M., ... Lüthi, A.


Viviani, D., Charlet, A., van den Burg, E., Robinet, C., Hurni, N., Abatis, M., . . . Stoop, R.


