7-18-2014 12:00 AM

The Neural Correlates of Guilt and Restitution During a Social Decision Making Task

Ambrose Ty
The University of Western Ontario

Supervisor
Elizabeth Finger
The University of Western Ontario

Graduate Program in Neuroscience
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

THE NEURAL CORRELATES OF GUILT AND RESTITUTION DURING A SOCIAL DECISION MAKING TASK

(Thesis format: Monograph)

by

Ambrose Ty

Graduate Program in Neuroscience

A thesis submitted in partial fulfillment
of the requirements for the degree of
Masters of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Guilt is a social emotion that promotes prosocial and moral behaviours. It arises as a result of harming another individual, serving as a prompt for the guilty individual to take reparative actions, known as restitution. The neural regions that are involved in guilt and restitution, however, are not currently known. To identify these regions, we employed a novel social decision-making fMRI paradigm involving donations to charities. There was a significant positive correlation between trait guilt and BOLD signal in the vlPFC and mPFC during acts of restitution. Furthermore, choices of harm when compared to help showed increased BOLD signal in the amygdala, insula, and the superior temporal sulcus. The present results are consistent with past studies that indicate an important role of the vlPFC and mPFC for processing aversive social cues and to resolve decision conflict.

Keywords

Guilt, Social decision making, fMRI, charities

Acknowledgements

I would like to thank and show my appreciation to my supervisor, Dr. Elizabeth Finger. I am truly thankful for the way you have provided guidance and have been so encouraging throughout these last two years from the start with the initial brainstorming of ideas for the thesis to the designing of the task to the data acquisition and analysis and finally, to this final product that is now my thesis. I would also like to thank Dr. Derek Mitchell and the Mitchell lab for all their input and feedback throughout the process of designing my study as well, it has been greatly appreciated and has helped me a lot along the way. Finally, I want to also thank my advisory committee, Dr. Paul Frewen and Dr. Stefan Everling for their input as well towards this thesis.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

vlPFC: Ventrolateral Prefrontal Cortex

dmPFC: Dorsomedial Prefrontal Cortex

dlPFC: dorsolateral Prefrontal Cortex

vmPFC: Ventromedial Prefrontal Cortex

SMA: Supplementary Motor Area

dACC: Dorsal Anterior Cingulate Cortex

TPJ: Temporal Parietal Junction

NAcc: Nucleus Accumbens

FTD: Frontotemporal Dementia

bvFTD: Behavioural Variant Frontotemporal Dementia

ASPD: Antisocial Personality Disorder

PPD: Psychopathic Personality Disorder

ToM: Theory of Mind

BA: Brodmann Area

ISI: Interstimulus Interval

IAPS: International Affective Picture System

PPI-R: Psychopathic Personality Index - Revised

TOSCA: Test of Self-Conscious Affect

AFNI: Analysis of Functional Neuroimages

Chapter 1 Introduction

1.1 Research in Morality and Social Emotions

Imagine yourself in a situation where a runaway trolley in headed for five railway workers and you had the opportunity to save these five workers by pulling a lever that would change course of the trolley and would kill one person instead. It is likely that you would find it appropriate to pull the lever. What if, instead of pulling the lever, you had to push that person onto the track in order to stop the trolley and save the five other people? The decision suddenly becomes much harder and you would likely choose not to push the person, despite the fact that the overall outcome of the decision remains the same. Now, what if you were a surgeon that had the decision to kill one patient and take their organs in order to transplant them and save five other patients? All of a sudden, the answer becomes obvious that you would not do such a thing. In all of these scenarios, there appears to be another aspect beyond a simple logical reasoning that influences your answer, something that we call morality.

Morality has been a topic of interest throughout history but was largely reserved for philosophy until more recently, where there has been an increasing interest to empirically study this concept. Current research has shown that there are two aspects to moral decision making: a cognitive and an affective aspect. The cognitive aspect refers to moral reasoning, maximizing good while reducing harm, and cognitive perspective taking or theory of mind, the ability to view situations from another person's perspective. The affective aspect refers to the social emotions of shame, guilt, empathy, and sympathy. Each of these emotions motivates moral behaviours through different manners, for example guilt occurs through a failure to meet one's moral standards whereas shame occurs due to failure of meeting other's moral standards. In both situations, it is a negative emotional response that drives an individual to promote moral behaviours, such as choosing not to push a person onto the train tracks to stop the train, even though it serves the aggregate welfare or the greater good. Although there have been more recent findings from imaging studies of healthy and of patient populations showing that while there are neural regions showing interplay between the two aspects, they are still dissociable from one another (Carlo *et al.* 2012). For example, a study by Greene (2001), showed that people who select the utilitarian response, the choice to push the person, have increased activation in areas of abstract reasoning and cognitive functions, such as the dorsolateral prefrontal cortex and anterior cingulate. On the other hand, people who choose not to push the individual show increased activation in the ventromedial prefrontal cortex and other areas of the brain that subserve emotions, such as the insula.

1.2 Guilt and Shame

The social emotions that govern the affective aspect of morality were once all thought to be one fundamental emotion. This was due to the fact that these emotions occur in response to very similar situations and are often confused with one another and resulted in them being used interchangeably. For example, guilt and shame are the two most commonly interchanged and embarrassment was simply thought to be a milder form of shame (Borg, Staufenbiel, & Scherer, 1988 & Tangney *et al.*, 1996). Tangney *et al.* (1996) sought to differentiate these emotions from one another and characterized these emotions on four themes, the situations that elicit the emotions, the feelings associated with the emotion, the cognition during the emotion, and the actions that are caused by these emotions. Using these four aspects, they concluded that these three emotions were uniquely different, not only in intensity, but in the way each are construed and experienced. Since then, these emotions have been characterized through numerous studies. Guilt and shame are still frequently confused, but the largest difference between these two emotions lies in their function (Tangney & Dearing, 2002 & Carlo *et al.* 2012).

Guilt has been described as the "social mortar" of human society (De Hooge *et al.*, 2011). It is involved in the evaluation of the self and occurs when behaviours or actions do not meet one's own moral standards, which ultimately motivates prosocial behaviours (Carlo *et al.*, 2012, Tangney & Dearing 2002, Haidt, 2003, Carni *et al.*, 2013). The goal of guilt is to motivate moral behaviour and has been found to be highly linked to altruism, the act of helping without any expectance of return (Haidt, 2003 & Carni, 2013). On the other hand, shame is a highly negative emotion that occurs when one fails to meet the standards of other people (Carlo *et al.* 2012). It is a painful experience that is associated feelings of exposure and belitting of onself, and ultimately promotes behaviours to protect one's self-image (Tangney *et al.* 1996, Carlo *et al.* 2012, Carni *et al.* 2013). Guilt and shame, therefore, are distinct emotions that serve separate functions and occur in response to different situations.

1.2.1 Guilt and Moral Behaviours

It should be apparent that guilt is a social emotion deeply engrossed into the moral realm, promoting moral behaviours in people. Berndsen *et al.* (2004) concluded that guilt arises from actions that lead to negative interpersonal consequences, and as the level of interpersonal harm increased, so did the feelings of guilt. The helpfulness of guilt in promoting social behaviours can be seen in a study by Ketelaar & Tung Au (2003) where participants were asked to describe a situation in which they felt guilty, ashamed, or self-blaming prior to playing either the prisoner's dilemma or the ultimatum game. In both cases, individuals who wrote a detailed description of the guilty story were more cooperative and were more generous when compared to the other two conditions. The types of behaviours guilt promotes and process of how it does so may still seem unclear and the literature remains sparse. Carlo *et al.* (2012), however, sought to determine the types of prosocial behaviours elicited by guilt and divided them into 6 types of moral categories: altruistic behaviours, public prosocial behaviours, compliant prosocial behaviours, anonymous prosocial behaviours, dire prosocial behaviours, and emotional prosocial behaviours. It was found that guilt was positively related to compliant behaviours, that is, helping those who request aid, altruistic behaviours, i.e. helping without expecting self-reward, dire prosocial behaviours (helping during crisis situations), and emotional helping, which is helping during "affectively evocative" contexts. In addition to this, guilt was found to be a stronger motivator of these behaviours than sympathy and was negatively related to public and aggressive behaviours. Thus it appears that guilt is important in promoting positive behaviours while inhibiting negative or "bad" behaviours (Carlo *et al.*, 2012 & Carni *et al.,* 2013). In addition to this, it is important to note that guilt specifically promotes these types of behaviours and is not due to a general sense of immorality (Cryder *et al.*, 2012). These studies have provided empirical evidence for observations made in the past. Individuals that are more prone to guilt have been found to have better interpersonal relationships, are less likely to react in anger to situations, and are more willing to accept wrongs that they have committed and is linked to promoting positive moral development (Tangey & Dearing, 2002 & Berndsen *et al.*, 2004 & Laible *et al.,* 2008). Guilt is still a high-order complex emotion that will still require much more stringent research in order to fully tease it apart and for us to understand it.

1.2.2 Physiological Responses of Guilt

The somatic theory of emotion states that we are able to identify emotions and process emotional information by monitoring our bodily responses and labelling it (Wahlund *et al.*, 2010). While this theory is still faced with much debate and may not be the sole method for identifying emotions, it provides a clue towards bodily changes that occur concurrently with emotional experiences. For example, individuals who wear heavier backpacks were found to show increased guilt, where the physical experience of weight was linked to the subjective "weight" of guilt. Individuals wearing heavier backpacks showed reduced likelihood of engaging in guilt-inducing behaviours when compared to those wearing lighter backpacks (Kouchaki *et al.* 2013). Furthermore, this avoidance of guilt-inducing behaviours has been found to occur even in physiological changes such as heart rate and skin conductance. A study by Oliveira-Silva & Goncalves (2011) investigated empathy, a pre-requisite and key component of guilt. In their study, emotional empathy was correlated with increased arousal as measured by increased skin conductance and cognitive empathy was correlated with increased heart rate. In another study, a test called the Guilty Actions Test was used. In the test, participants committed or witnessed a mock crime and then were asked to respond to questions that were related to the crime. When individuals read answers related to the crime, it was found that skin conductance was increased (Gamer *et al.* 2008). Wahlund *et al.* (2010) furthered this conclusion by studying physiological changes in individuals with antisocial personality disorder. These individuals have long been known to have reduced emotional experience, but it was found that this was also reflected in attenuated autonomic arousal as well. Therefore, physiological changes that are congruent with and accompany emotions are important for and can increase emotional experience (Niedenthal *et al.* 2007, Gamer *et al.*, 2008, Oliveira-Silva & Goncalves, 2011, Kouchaki *et al.*, 2013).

1.3 Anticipatory Guilt

While emotions themselves are able to strongly affect behaviours, they occur in response to situations that have already occurred. In the case of guilt, the prosocial behavioural change evoked by guilt occurs after the fact of the harming or suffering of an individual. This thought appears to be troubling, if harm necessarily needs to precede prosocial behaviours. However, there is a concept of emotional avoidance, where the negative experiences of an emotion act as a deterrent to employ behaviours that may result in re-experience (Wagner *et al.* 2012). In the case of guilt, we will refer to this concept as anticipatory guilt. The literature on the concept of anticipatory guilt is scarce, thus further research needs to be done to better understand it (Lindsey *et al.*, 2005).

Most of the studies on anticipatory guilt revolve around organ donation, where individuals are asked to decide whether they are willing to sign an organ donor card after reading a description about the benefits of donating (Lindsey *et al.,* 2005, Lindsey *et al.*, 2007, Wang, *2011*). In these studies, the presentation of the need of unknown others of donation vary and certain guilt inducing sentences or phrases are included into the description. Lindsey *et al.* (2005) found that individuals who predicted higher feelings of guilt if they did not help were more likely engage in helping 7-10 days later. Additionally, these individuals that do help felt less guilty than those that do not help. In a follow up study, Lindsey *et al.* (2007) studied whether empathy, the precedent for guilt, is related to anticipatory guilt. This is due to the fact that there is no apparent reason to alleviate this threat of death (to the unknown other) that has no effect on the individual, and thus it is of interest to understand whether empathy is related to anticipatory guilt. Although past research suggests that thought alone can induce anticipatory guilt and cause avoidance of behaviours that would elicit guilt, this study found no correlation (O'Keefe, 2002). Thus, while anticipatory guilt is able to predict organ donation and other prosocial behaviours, further research needs to be done with regards to empathy. Wang (2011) confirmed the findings that anticipatory guilt increase helping behaviours and does so by influencing attitudes and norms on behavioural intentions. Also, it was suggested that anticipatory guilt is more influential on behaviours that have a higher and more direct impact on others. In summary, anticipatory guilt occurs from the speculation of likely emotional consequences from future behaviours, which drives prosocial behaviours by shaping and altering behavioural intentions and ultimately lowers feelings of guilt.

1.3.1 Neuroimaging data on Anticipatory Guilt

To date, there is only one study that has directly looked at the neural correlates during the avoidance of guilt. Chang *et al.* (2011) looked at guilt avoidance with an economical investment game. Participants were asked to invest money to another person, whom would then receive four times that amount, and decide how much they would expect in return. i.e. A donates 10 dollars to B, who receives $40, and A expects a $20 return. In the second part of the study, a subset of the original participants played the role of B, while inside an fMRI scanner. The results showed that player B was accurate at estimating the amount of money player A was expecting, and thus avoided guilt. When this occurred, the insula, supplementary motor area (SMA), dorsal anterior cingulate cortex (dACC), dorsolateral prefrontal cortex (dlPFC), and temporal parietal junction (TPJ) were activated. When the estimation was lower than the expectation (increasing selfish financial gain), the ventromedial prefrontal cortex (vmPFC), bilateral nucleus accumbens (NAcc), and dorsomedial prefrontal cortex (dmPFC) were activated. These areas are similar to those of neuroimaging data found in guilt, and thus of particular interest are the overlapping areas: the insula, SMA, dACC, dlPFC, and TPJ.

1.4 Deficits of Guilt in Neuropsychiatric Disorders

Neuropsychiatric disorders that exhibit deficits in guilt have such as frontotemporal dementia (FTD) and antisocial personality disorder (ASPD) have provided another perspective towards understanding the function of guilt. However, determining the neural basis of guilt in healthy individuals may also serve to understand these neuropsychiatric disorders better.

1.4.1 Frontotemporal Dementia

Frontotemporal dementia (FTD) is a unique form of dementia that is characterized by atrophy of the frontal and temporal lobes. In addition to this, another unique hallmark of FTD patients is progressive and drastic social and personality changes (Mendez *et al.*, 2005, Mendez *et al.,* 2006). These patients are often found to perform socially inappropriate behaviours that border sociopathic behaviours such as unsolicited sexual behaviours, traffic violations, physical assaults, stealing, and breaking and entering (Miller *et al.* 1997, Mendez *et al.*, 2005, & Mendez *et al.* 2006). These changes are particularly pronounced, especially early at the onset of the disease, in one of the subtypes of FTD, behavioural variant FTD (bvFTD), where the atrophy is focalized to the frontal lobes. The behavioural changes have been studied extensively and have been found to stem from a blunting of emotions, loss of empathy, and clinical apathy, where emotional blunting refers to things such as emotional shallowness, a lack of concern for others, and loss of emotional warmth (Mendez *et al.*, 2006). These deficits in normal aspects of social functioning have been found to be correlated with the degree of atrophy. More specifically, these areas of atrophy include frontal regions such as the vmPFC, dlPFC, OFC, the anterior insula, and anterior cingulate and other areas that have been implicated in theory of mind (ToM) such as the temporal parietal junction and the superior temporal gyrus (Eslinger & Moore, 2011, Eslinger *et al.*, 2012, Mendez & Shapira, 2011, & Moll *et al.* 2011). An interesting aspect of this emotional blunting and change in behaviour is the fact that these individuals are still able logically reason out socially appropriate answers when not limited by time (Moll *et al.*, 2005). In a study where FTD patients were asked to make a decision in moral dilemmas, similar to the ones that were referred to at the beginning of the introduction, they almost always chose the utilitarian answer, that is to sacrifice the lesser for the greater, regardless of the scenario (Moll *et al.*, 2005). The investigators classified the patients' reasoning as "impersonal" and attribute this to a lack of emotionality in their reasoning that arises from damage to the vmPFC. It appears that FTD patients still have moral and social knowledge and rules and are able to consciously identify them, however, this loss of emotional "colouring" to emphasize certain stimuli is the major reason for the inappropriate behaviours.

1.4.2 Antisocial Personality Disorder (ASPD)

ASPD is a personality disorder that is characterized by the core traits of callous and unemotional personality, including characteristics of lack of empathy or remorse, uncaring nature, or shallow emotional responding (Herpertz *et al.*, 2001; Cooke *et al.*, 2005; Frick & White, 2008; Marsh *et al.*, 2013). The clinical population of ASPD, however, can be split into two sub-populations based upon behavioural patterns, although we will focus only upon of these sub-populations. Approximately 30% of the population show a lack of empathy and response, instrumental aggression, and deficient affective experience, including guilt (Frick & White, 2008 & Gregory *et al.*, 2012). This subpopulation has been referred to as psychopathic personality disorder (PPD) and these individuals show decreased electrodermal response to punishment or anxiety-related stimuli, though to indicate low levels of fear. Additionally, these individuals show an absence of startle potentiation, which refers to lack of increased startle response when a loud acoustic probe is presented with a negative visual stimuli. In addition to this lack of fear, studies have looked this clinical population and found that these individuals have lowered responding to emotional slides and a lack of modulation of startle response by any emotional stimulus (Herpertz *et al.*, 2001). This hypoemotionality is thought to predispose these individuals to antisocial and inappropriate behaviours (Herpertz *et al.*, 2001). These behaviours have been suggested to be linked to the frontal cortices, more specifically the vmPFC and the orbitofrontal cortices. Early damage to the vmPFC has been found to result in behaviours similar to that of sociopathy and PPD, most importantly showing insensitivity to future consequences of decisions (Anderson *et al.*, 1999). Similarly, Blair & Cipolotti (2000) found similar behavioural changes in an adult suffering damage to the frontal lobes, including the orbitofrontal cortices. As well, children that show psychopathic traits were shown to have atypical vmPFC activation during a reversal learning task, specifically for punished errors (Finger *et al.* 2008). The guilt deficits found in persons with PPD could be due to abnormalities in vmPFC and orbitofrontal cortex function, which are important for appropriate social behaviours, empathy, and appropriate consideration of future consequences of decisions. Elucidation of the neural regions and functions supporting social decision making and guilt would be expected to provide insight into the neural basis of guilt and related deficits in neuropsychiatric disorders.

1.5 Neurocorrelates of Guilt

While the above patient studies may help to identify the neural regions supporting guilt-related processing, more recent work has attempted to determine the areas activated in healthy populations during the generation of guilt and related social decision-making. Convergent with the patient findings described above, these studies have identified specific regions of prefrontal cortex, the amygdala and temporal lobes during such tasks. Two fMRI studies have evaluated guilt and embarrassment by presenting sentences to participants designed to elicit moral emotions. Takahashi *et al.* (2004) found that social emotions of guilt and embarrassment produced activation within neural regions implicated in theory of mind (ToM) processes, the medial prefrontal cortex, left posterior superior temporal sulcus, and the visual cortex. Elicitation of guilt was associated with higher activation in the medial prefrontal cortex (mPFC), a region implicated in monitoring mental states and making moral judgements (Adolphs, 2001; Greene & Haidt, 2002; Green *et al.* 2001). Finger *et al.* (2006) expanded upon these findings by introducing the presence or absence of an audience during scripts read to elicit guilt or embarrassment. In doing so, this allowed one to identify regions that may prompt behavioural change following a social or moral transgression; where moral transgressions refer to transgressions that are considered inappropriate or wrong regardless of the presence of a witness or not (i.e. murder) and social transgressions are wrong only when witnessed (i.e. burping). They found that the ventrolateral prefrontal cortex (vlPFC) and dorsomedial prefrontal cortex (dmPFC) showed greater activation regardless of audience presence during moral transgressions but only in the presence of an audience for social transgressions. This finding was interpreted in the context of previous research that has found that the vlPFC and the dmPFC are important for processing aversive social cues and to resolve conflict between incompatible voluntary action plans (Blair *et al.* 1999). As mentioned previously, Chang *et al.* (2011) studied guilt aversion, to determine whether anticipation of guilt motivates prosocial behaviours. In contrast to the previous studies mentioned above, when making decisions that should minimize guilt, increased activation was found in the insula, supplementary motor area (SMA), dorsal anterior cingulate (dACC), dorsolateral prefrontal cortex (dlPFC), and the temporal parietal junction (TPJ).

These studies indicate a role for dorsomedial and ventrolateral prefrontal cortex, the superior temporal sulcus, the insula, anterior cingulate cortex and amygdala in guilt and related decision making. A brief review of these areas will now be presented to highlight the general cognitive processes currently attributed to each, as well as their more specific role related to decision making and moral emotions, specifically guilt.

1.5.1 Dorsomedial Prefrontal Cortex

The dorsomedial prefrontal cortex (dmPFC) refers to the medial portions of Broddman Areas (BA) 8 and 9. The dmPFC has been found to play a role in error monitoring as well as learning and reward processing in complex decision making, leading to adjustments in behaviour to optimize reward (Venkatraman *et al.* 2012). The dmPFC is activated during the Stroop task and its variants, where there is conflict between a prepotent response and the task response and participants must choose the latter behaviour over the former (Bush *et al.*, 1998; Derrfuss *et al.,* 2005). In addition to this, the activity in the dmPFC has also been shown to be correlated with increasing decision conflict. In a study by Pochon and colleagues, male participants were asked to pick between two equally attractive female faces or between one attractive and one unattractive female face. When asked to choose between two equally attractive faces, that is when there is a greater conflict, the dmPFC showed greater activation when compared to the low conflict condition (attractive vs unattractive face). The dmPFC is also found to be active when actions lead to future errors and my predict error likelihoods even before receiving feedback (Ullsperger and von Cramon, 2003; Brown and Braver, 2005). Additionally, Kennerley and colleagues (2006) lesioned the dmPFC in macaques and found that they became unable to use past reward informations to guide future behaviours (Kennerly *et al.*, 2006). These suggest that the dmPFC integrates information about past actions and outcomes and is used to guide future behaviours. This is also supported by studies using magnetic resonance tractography and functional connectivity which have revealed connections between dmPFC and areas involved in multimodal sensory representation, visuospatial processes, and memory (Beckmann *et al.* 2009; Venkatraman *et al*., 2009b). This puts the dmPFC in a position to modulate various processes in the brain to alter behaviour in response to new information (Mitchell *et al.*, 2013). Additionally, the dmPFC has strong connections to the dorsolateral prefrontal cortex, an area important for making choices between multiple responses that are contingent upon the context (Miller and Cohen, 2011; Taren *et al.*, 2011). Thus, the dmPFC is an area of particular interest in decision making due to its role in selecting or biasing certain behaviours based upon past errors and knowledge.

1.5.2 Ventrolateral Prefrontal Cortex

The ventrolateral prefrontal cortex in this current study refers to Broddman Area 47. Similar to the dmPFC, it has many anatomical connections with various sensory modalities in order to influence behaviour, and most notably, with the anterior insula, an area that has been implicated in emotions and decision making (Singer *et al.*, 2009). The vlPFC has been found to be activated in reversal learning, particularly in response to punishment cues (Elliot *et al.*, 2000; Finger Arch Gen Psychiatry, 2007). Cools *et al.* (2002) found that activity in the vlPFC occurred only in response to errors that prompted subsequent reversal behavioural changes. Additionally, this activity has been shown to occur independent of punishment, but rather it is specifically responding to sub-optimal behaviours. Mitchell *et al.* (2009) employed an object discrimination task, in which participants had to pick between pairs of images of objects that had a pre-assigned point value and were asked to optimize the gain of points, as the point value of the objects changed over the course of the study. The vlPFC was only found to activate in response to sub-optimal responses, that is when there was a reversal error when participants chose an object in a pair that now had a lesser value. The vlPFC has been suggested to play a role in processing context relevant stimuli and in modulating behaviours (Mitchell *et al.*, 2013).

In summary, the dorsomedial and ventrolateral prefrontal cortex have been implicated in many cognitive functions, ranging from executive functions such as decision making to emotion regulation. In conjunction with these and their findings, Finger *et al.* (2007) hypothesized that these areas may activate in response to aversive cues to facilitate a behavioural change. In that study, however, the participants were passively reading the sentences and there was no opportunity for active behavioural change in the task. Thus, in the current study, we plan to determine whether dmPFC and vlPFC may activate in response to decisions that elicit guilt, specifically to prompt or facilitate restitutive actions towards the victim.

1.5.3 Insula

The insula is an area of the brain that is highly interconnected to many subcortical and cortical areas. Of particular interest to the current study is the anterior portion of the insula, which has been found to be important for interoceptive representations of the body and emotions

(Shin *et al.*, 2000). The insula has also been implicated in negative affective states such as in guilt (Shin et al., 2000) and in disgust (Calder *et al.*, 2000; Wicker *et al.*, 2003). Specifically in disgust, the anterior insula was found to play a role in both the experience of the emotion and in the perceiving and understanding of the same emotion in others, an important aspect of empathy (Wicker, 2003). In support of this, the anterior insula is interconnected with limbic structures such as the amygdala, an area implicated in emotional processing, especially of negative emotions (Flynn, 1999).

More recently, it was hypothesized that the insula may reflect the subjective experience of emotions and the processing of predicting and representing future uncertainties. In the study by Chang *et al.* (2011), they found that when participants minimized future feelings of guilt, by avoiding guilt inducing behaviour in a Trust Game, the insula showed increased activity and the resultant pattern of activation within the brain was similar to making decisions of rejecting unfair offers in the Ultimatum Game (Sanfey *et al.*, 2003). Ullsperger (2003) found that the anterior inferior insula showed greatest activity when receiving negative feedback during a motion discrimination task, suggesting a role in responding to negative emotional states elicited by the feedback. In conjunction with this, a prior study in our lab found that activity within the anterior insula was associated with decisions and feedback in a charity task (Greening *et al.*, 2013). In the task, participants were asked to assign losses or gains to either themselves or a charity; however, at one-third chance, these assignments of gains or losses would be flipped. That is, there is a one in three chance that an assignment of gain could result in a loss. During these unintended outcomes compared to intended outcomes, the anterior insula showed greatest activity. Furthermore, increase anterior insula activity was associated with a greater tendency to avoid harming the charity (avoid taking money from the charity). Together these findings suggest a role in calculation or anticipation of risk, which can influence decision making. Thus it appears that the insula has a potential role in emotional decision making, related to avoiding future negative feedback and potentially specific emotional states, such as guilt.

1.5.4 Amygdala

The amygdala is commonly known as the "fear centre" of the brain, important for detecting danger in the environment, and has been studied extensively in this regard in nonhuman animals (Adolphs *et al.*, 1999; Sergerie *et al.*, 2007; Todd & Anderson, 2009). More recently, however, this view has begun to change in light of new studies in humans. The role of the amygdala has been found to not only process negative but also positive emotional stimuli (Sergerie *et al.*, 2007 & Anderson, 2009). Emotions act as a guidance cue to biologically relevant information, as shown through the attentional blink paradigm. In this paradigm, there is often an attentional "blink" where if the two targets are presented temporally proximal enough, the second target is often missed. However, if the target has some sort of emotional content, it is less likely to suffer from this attentional blink. In patients with amygdala lesions, this benefit of emotional content was not found and therefore, affective significance of stimuli may determine the fate of visual stimuli (Anderson & Phelps, 2001).

As a result, the amygdala has been hypothesized to have a role in the detection of biological and social relevant information, a "relevance detector" (Sander *et al.*, 2003 & Sergerie *et al.*, 2007). Connectivity of the amygdala to various areas within the brain, particularly the visual areas, reveals a potential mechanism for a modulatory effect on the processing of emotionally-laden stimuli (Schwartz *et al.*, 2002; Cardinal *et al.,* 2002; Sergerie *et al.,* 2007; Heimer *et al.,* 2008). Individuals suffering from Kluver-Bucy syndrome with specific bilateral amygdala lesion, which has shed much light in this regard. SM, one of these patients, fails to identify facial expressions of emotion such as surprise, fear, and anger (Adolphs *et al.* 1995), and follow up studies have found that this impairment is for all unpleasant emotions in general (Adolphs *et al.* 1999a). In a study where individuals that suffer from amygdala damage were compared to controls in an emotion discrimination task, it was found that controls have higher amygdala response to fearful faces than neutral whereas individuals with the amygdala damage did not (Adolphs *et al.*, 1999b). Thus the amygdala appears to support social networks within the brain, thereby influencing the way by which we interact with the social world around us by modulating salience based upon emotional content.

1.5.5 Superior Temporal Sulcus (Temporoparietal Junction)

The superior temporal sulcus has been implicated to be part of a more general social cognition network that is activated during processes that involve other intentional human agents, in tasks such as the prisoner's dilemma game, theory of mental states tasks, and dispositional attribution tasks (Singer *et al.*, 2004, Harris *et al.*, 2005, Lee & Harris, 2013). This network has also found to be show activation in previous studies that have explored the identification of moral or social transgressions (Greene *et al.*, 2001; Takahashi *et al.*, 2004; Finger *et al.*, 2006). Particularly, Finger *et al.* (2006) found that the STS demonstrated significantly greater activity during moral transgressions, which were designed to elicit guilt in participants.

1.6 Thesis Hypothesis

While many studies have shown that guilt results in more altruistic and helping behaviour (Lindsey, 2005; Lindsey & Hill, 2007; Harbaugh *et al.*, 2007), the neural regions supporting the emotion of guilt remain to be confirmed, and those supporting acts of restitution have not been previously evaluated. In many of the prior studies mentioned above, the participants passively read words, sentences, or paragraphs from a screen during fMRI. Therefore, it is also of interest to determine whether agency has an effect upon the neural structures implicated in guilt generation, given the known links between degree of agency and intensity of guilt feelings

(Berndsen *et al.*, 2004). In the current study we will test the hypotheses that 1) dmPFC and vlPFC activity during social decision making which harms another reflects processes that underlie acts of restitution, and 2) that anticipated and experienced guilt are represented by overlapping networks involving the anterior insula, amygdala and temporoparietal junction. We will test our hypothesis by examining neural activity using fMRI during a novel social decision-making task designed to produce guilt and offer opportunities for and restitution.

Chapter 2 Methods

2.1 Participants

Participants were recruited through posters that were placed on The University of Western Ontario campus and were also directly contacted from a participant research pool that have participated in previous studies. Participants were screened prior to the study for contraindications to MRI and for handedness. Study exclusion criteria included recent neuropsychiatric disorders or current medication for the disorder. In total, 23 healthy participants were recruited for the study. However, neuroimaging data was available for 18 healthy participants; data was unavailable from 5 participants due to scanner dysfunction or a lack of response during the scan. In total, there were 9 males and 9 females, all participants were right handed, with an average age = 20.28 years (Range: 18-23; S.D.: 1.52). Participants responded to 91% of the 80 trials (Mean = 72.95; S.D.: 2.36; Range: 56-80). All participants were provided written informed consent. The procedures were approved by The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB protocol #13617).

2.2 fMRI task

2.2.1 Study Overview

The participants performed a novel donation task, where they were given the opportunity to donate to individuals of need. Healthy volunteers were contacted and invited to participate in a "Social Decision Making Study". Upon arrival, participants read a set of instructions that explained the task paradigm. They were then presented with a list of fictional charities on a website, to enhance the believability of the study. They were told that the money would go to these charities and would help individuals in need; however, they were also told that a portion of money that they choose not to donate would be added as a bonus towards their own compensation for the study. This was to increase the tension between self-rewarding vs. prosocial behaviours that is typical of real-world altruistic decisions, and to further intensify the corresponding emotions such as guilt that may be evoked by their decisions. The agency of participants in making decisions is of interest here due to the positive correlation between agency and intensity of feelings of guilt. Following study completion, participants were debriefed on the fictitious nature of the charities.

2.2.2 fMRI Task Design

The task paradigm (Fig 1) consisted of guilt-inducing scenarios with two conditions, active or passive, that were presented to the participant. In both conditions, participants read about an individual in need of help from a charity that was paired with a negative image of a person from the International Affective Picture System (IAPS). In the active condition, participants were given a choice to donate $10 or not donate to that particular charity. A feedback screen then appeared, depending on their response, that was designed to either induce guilt (further harm to the victim described after a no donation choice) or was neutral (following a donation). The participants were then given a second chance to donate. Following our piloting of the task, to increase the number of trials where a donation decision was made during choice 2, on 50% of the trials, the amount to be donated in choice 2 was halved to $5, while for 50% of trials the amount remained $10. A second feedback screen appeared after choice 2. In the passive condition, donation choices were randomly selected by the computer, so that the number of donate and no donate selections was balanced. Participants simply viewed the choices and were asked to read the feedback (Appendix 1). Thus, in each trial there were two decisions to donate or not donate, resulting in four total possible combinations of responses, where donations are “help” and no donations are “harm”: Help/Help, Help/Harm, Harm/Help, and Harm/Harm.

In total, there were 80 unique scenarios that described the situations of individuals that were in need. All scenarios were balanced in both character and word length. Each scenario was presented once in the active condition and once in the passive condition, thus totalling 160 trials altogether. Each trial consisted of the presentation of the individual in need two choice screens and two feedback screens, together lasting 13.5 seconds Two variable interstimulus intervals of either 0.5, 1.5, or 2.5 seconds were placed between choice and feedback screens to permit dissociation of the BOLD responses between these two events. The variable ISI was programmed to equally select between the three intervals, averaging a total of 1.5 seconds and thus on average, each trial lasted 16.5 seconds. The 160 trials were split into 6 separate runs, with two runs lasting a total of 429 seconds or 7 minutes and 9 seconds and four runs lasting a total of 445.5 or 7 minutes and 25.5 seconds. For each participant, the run order was randomly selected and within each run, the scenarios were also presented in a random order.

2.2.3 Post-Task Emotional Ratings Task

In order to determine whether the scenarios and associated decisions aroused the anticipated emotions in participants, and to permit correlation of BOLD signal with subjective emotion ratings, after the scan, individuals completed a second computer task where they were presented with the decisions they made for each trial during the scan and were subsequently asked to rate how compassionate or guilty they felt following each decision. Finally, three trait scales were also completed by participants to quantify their guilt-proneness and related traits, and to examine potential correlations between trait scores and BOLD responses during the task. The scales included the Test of Self Conscious Affect (TOSCA, Tangney *et al.*, 1989), Psychopathic Personality Index – Revised (PPI-R, Lilienfeld & Widows, 2005), and the Guilt Inventory (Jones *et al.*, 2000). The TOSCA is a questionnaire consisting of written responses to a series of brief scenarios that may elicit shame, guilt, and pride. It is used primarily to yield indices of shame, guilt, externalization, detachment/unconcern, alpha pride, and beta pride. The PPI-R is a self-report questionnaire that asks participants how likely they would engage in a certain behavior. It is used to assess three main areas of psychopathic personality: fearless dominance, impulsive antisociality, and coldheartedness. The Guilt Inventory is a scale that measures trait guilt, state guilt, and moral standards.

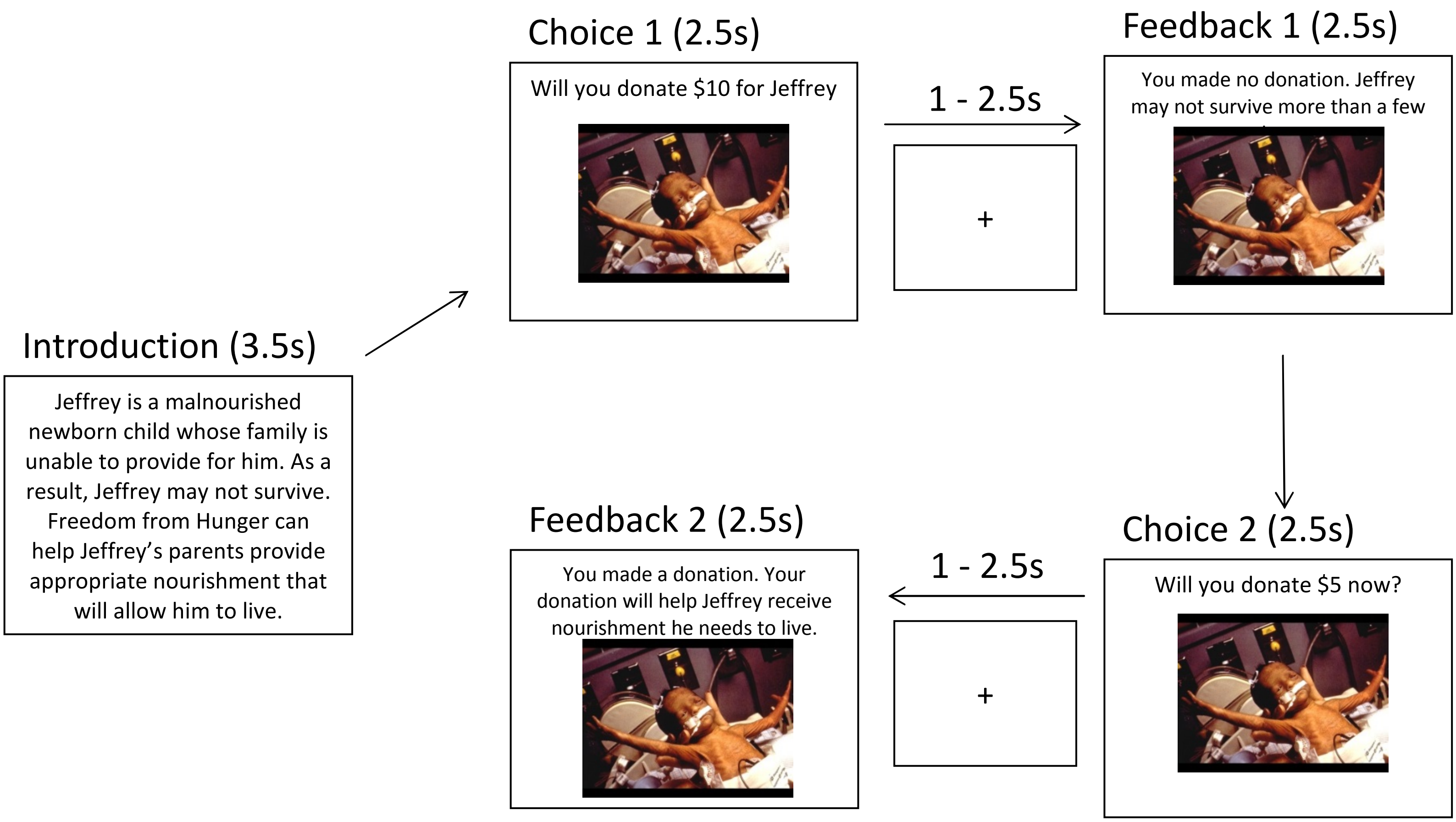

Figure 1. Example of Task Trial

2.3 Image Acquisition

Data was acquired using a 3-Tesla Siemens Tim Trio MRI scanner with a 32 channel head-coil at the Robarts Research Institute. Prior to the task, high-resolution structural-T1-weighted acquisition of anatomical data of complete brain volume was obtained in the axial plane: repetition time (TR) 2300ms, echo time (TE) 2.98 ms, field of view (FOV) 25.6 cm, 192 slices, voxel size = $1mm^3$, 256 x 256 matrix. Functional data was acquired using a gradient echo planar imaging sequence: 42 contiguous slices of 2 x 2 mm in plane, slice thickness 2.5 mm, TR 2500ms, TE 30 ms, FOV 24cm, 120 x 120 matrix.

2.4 Image Processing and Analysis

The fMRI data were preprocessed and analyzed using Analysis of Functional Neuroimages (AFNI; Cox, 1996) software package. The first 5 volumes of each run were discarded to ensure that magnetization equilibrium was reached. The functional images were motion corrected by registering all BOLD data to the first volume of the third run. Subjects' functional data were spatially smoothed with a 4 mm full-width at half-maximum (FWHM) Gaussian filter. The time series data were normalized by dividing the signal intensity of a voxel at each time point by the mean signal intensity of that voxel for each run and multiply the result by 100. Resultant regression coefficients represented the percent signal change relative to the mean. Regressors were created, characterizing trial condition (active or passive) and the donation response of the subject in response to each opportunity to donate (donate or no donate at choice 1 or choice 2) and also the feedback subsequent to the donation choice. In total 48 regressors were created by convolving the train of stimulus events with a gamma-variate hemodynamic response function to account for the slow hemodynamic response. Linear regression modelling was performed using the 48 regressors to model a first order baseline drift function, producing a beta coefficient and *t*-statistic for each voxel and regressor.

Individual anatomical and functional images were normalized to the standardized space of Talaraich and Tournoux. The group analysis was then performed through *t*-tests on whole brain data. The threshold was set at $p<0.005$ (corrected at $p < 0.05$ for multiple comparisons). Threshold correction was done using the 3dClustSim program in AFNI, which applied 10,000 Monte Carlo simulation iterations on a whole brain EPI matrix. This program calculates the number of adjacent clusters required at a specified threshold to determine the likelihood of significance.

Finally, in order to determine whether the BOLD signal in the functional data correlated significantly with subjective feelings of guilt, an amplitude modulated analysis was performed. The analysis works by taking the subjects' ratings of feelings of guilt towards each scenario and determines voxels that show activation significantly modulated proportionally by this behavioural measure of interest. This allowed for us to determine whether any brain regions showed a direct correlation between subjective feelings of guilt and BOLD signal.

2.5 Amplitude Modulated Analysis

The amplitude modulated analysis takes participants post-hoc trial-by-trial subjective ratings of feelings of guilt towards each scenario in the task and identifies voxels that show BOLD signal changes that are significantly modulated proportional to this behavioural measure of interest. This allowed for us to determine which brain regions may demonstrate activity most closely correlated with subjective feelings of guilt.

Chapter 3 Results

3.1 Trait Scales

Analysis of participants' scores on the three trait scales that were employed to measure levels of guilt proneness and callous and unemotional traits can be found in the tables below. Participant scores were not significantly different from standardized scores of similar population (Lilienfeld & Widows, 2005, See Appendix B).

Table 1. Average participant scores on the trait scales employed. TOSCA is the Test of Self-Conscious Affect and the PPI-R is the Psychopathic Personality Index-Revised

Guilt Inventory	Mean	SD	Low	High	Range
State Guilt	34	6.18	13	50	37
Moral Standards	44.18	8.33	23	68	45
Trait Guilt	54.24	12.53	25	97	72

TOSCA	Mean	SD	Low	High	Range
Shame	43.82	6.56	35	57	22
Guilt	57.4	8.07	47	75	28
Detached	28.12	6.63	21	41	20
Externalization	23.12	6.82	14	35	21
Alpha Pride	17.06	6.82	15	25	10
Beta Pride	17.82	3.26	15	25	10

PPI-R	Mean	SD	Low	High	Range
Machiavellian Egocentricity	41.18	9.26	26	58	32
Rebellious Nonconformity	32.06	6.98	20	45	25
Blame Externalization	25.35	4.43	21	36	15
Carefree Nonplanfulness	33.88	7.07	28	58	30
Social Influence	41.71	10.52	25	61	36
Fearlessness	30.82	9.69	14	50	36
Stress Immunity	30.00	7.29	18	44	26
Coldheartedness	30.24	8.02	21	47	26
Virtuous Responding	22.94	5.07	14	31	17
Deviant Responding	12.59	2.53	10	18	8
Total	265.24	37.99	202	344	142
Factor Self-Centred Impulsivity	132.47	19.43	110	174	64
Factor Fearless Dominance	102.53	21.68	64	145	81
Inconsistent Responding 15	8.12	2.63	4	15	11
SubIR40	15.24	3.92	10	23	13
Inconsistent Responding 40	23.35	5.62	14	35	21

3.2 Behavioural Data

3.2.1 Participant decision types

Participants' decision frequencies to donate or not donate during the fMRI task are shown below (Table 2). Help/Harm decisions had the highest amount of choices, whereas restitution had the least amount of choices. In general, event numbers across the 6 runs of the task were in range of the desired number for event related fMRI analysis (>=20) with the exceptions of the restitution condition (no donate/donate) and the donate/donate conditions. The Choice1 category shows the mean number of choices to help or harm during the first choice phase whereas Choice2 shows the average amount of choices made in each of the four possible outcomes.

Table 2. The average number of harm and help decisions for choice 1 and choice 2 made by participants over the course of the fMRI task

Behavioural Data						
	Choice 1		Choice 2			
	Help	Harm	Help/Help	Help/Harm	Harm/Help	Harm/Harm
Mean	42.06	29.76	13.76	26.41	9.59	18.94
SD	3.34	3.26	2.85	2.51	2.24	1.86

3.2.2 Trial by trial emotion ratings

To examine how different trials and decisions prompted social emotions, subjective ratings of guilt or compassion obtained from the post-hoc task were averaged across trial type (i.e. harm1, harm2, etc.) for each participant (Figure 2). T-tests were then conducted on the mean guilt and compassion ratings comparing guilt after the first decision and the second decision. In general, after participants chose to help in the second choice, subjective ratings of guilt significantly decreased, whereas after choices to harm, subjective ratings of guilt significantly increased.

Specifically, mean ratings of guilt decreased from 2.40 +/- 1.28 to 2.04 +/- 1.23 ($p<0.005$) after restitution (Harm-Help) and from 2.45 +/- 1.18 to 2.35 +/- 1.17 ($p < 0.05$) after

Help-Help responses. In the choices to harm, guilt significantly increased from 2.44 +/- 1.19 to 2.65 +/- 1.15, p<0.005 after Help-Harm choices and from 2.13 +/- 1.18 to 2.28 +/- 1.29 (p<0.005) after Harm-Harm choices (Figure 2). Feelings of compassion, however, had the opposite trend, increasing after harming decisions and decreasing after helpings decisions: the Help/Harm condition ratings of compassion decreased from 3.52 +/- 0.98 to 3.13 +/- 1.02 (p<0.005), the Harm/Harm decreased from 2.55 +/- 1.11 to 2.47 +/- 1.08 (p<0.05), while the Harm/Help increased from 2.78 +/- 1.09 to 3.14 +/- 1.06 (p<0.005).

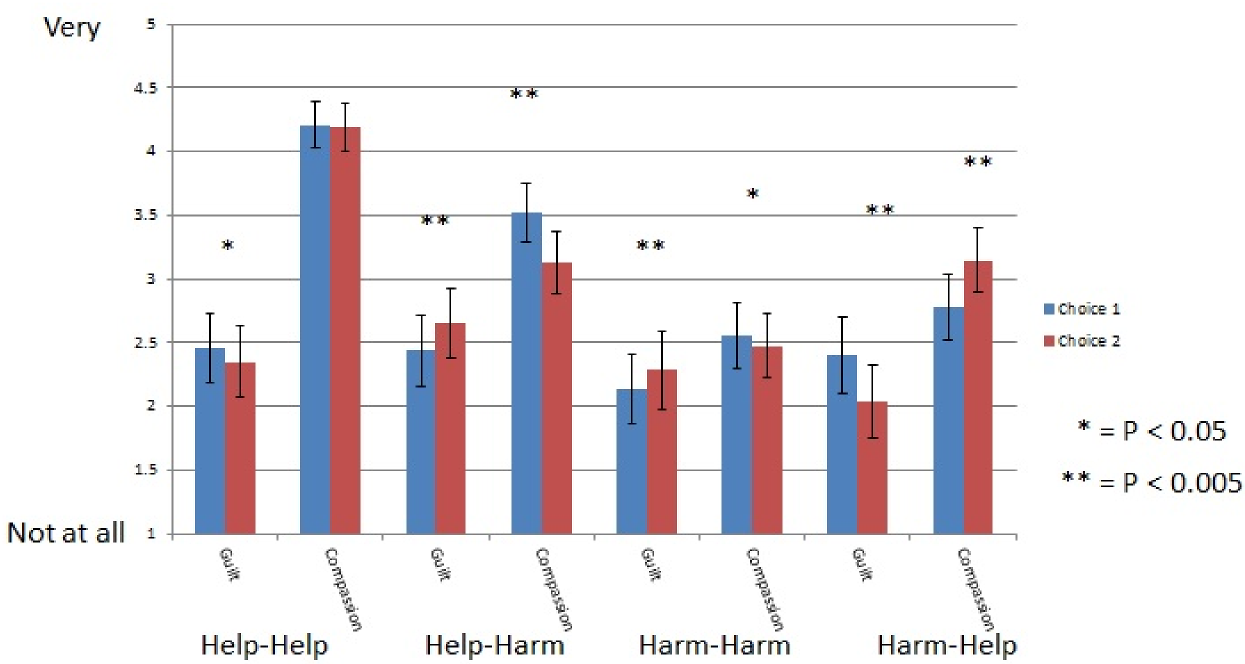

Figure 2. Subjective ratings of feelings of guilt and compassion collapsed into the four possible combinations of choices to harm or help individuals in need

3.2.3 Believability of Charities and Task Premise

Examination of participant ratings of the believability of the charities ("Did you consider these to be real charities?") demonstrated an average score of 3.5 (S.D. 0.3) out of 5, indicating general validity of the task premise (Figure 3). Examination of normality using the Shapiro-Wilk test of normality, shows that the distribution of ratings of considering the charities to be real was not normal D(18)=.86, p=0.01). However, examination of the response histogram did not demonstrate a bimodal distribution. Similarly, examination of participants' rating of the confidence that their donations to charities would help those in need ("To what extent were you confident that your donations would help the charities?") demonstrated a mean rating of 3.06 (S.D. 0.26) (Figure 3). The Shapiro-Wilk test indicated responses while responses for the confidence that donations did not significantly deviate from normality (D(18)=.90, p>.05).

3.3 Imaging Data

3.3.1 Elucidating the neural regions supporting the anticipation and experience of guilt

To determine the neural regions potentially supporting anticipation of guilt, we conducted a t-test on BOLD responses during the interstimulus interval (ISI) that followed a no-donate decision vs. the ISI following a donate decision, prior to receiving feedback. This demonstrated significantly greater activity following a donation compared to a non-donation in left sided regions including the superior, middle and frontal gyri and the middle temporal gyrus. Contrary to past research reporting increased activation in the insula during anticipation of guilt, this contrast, which we hypothesized would reflect anticipation of guilt, showed no brain regions demonstrating greater activation following a harm decision.

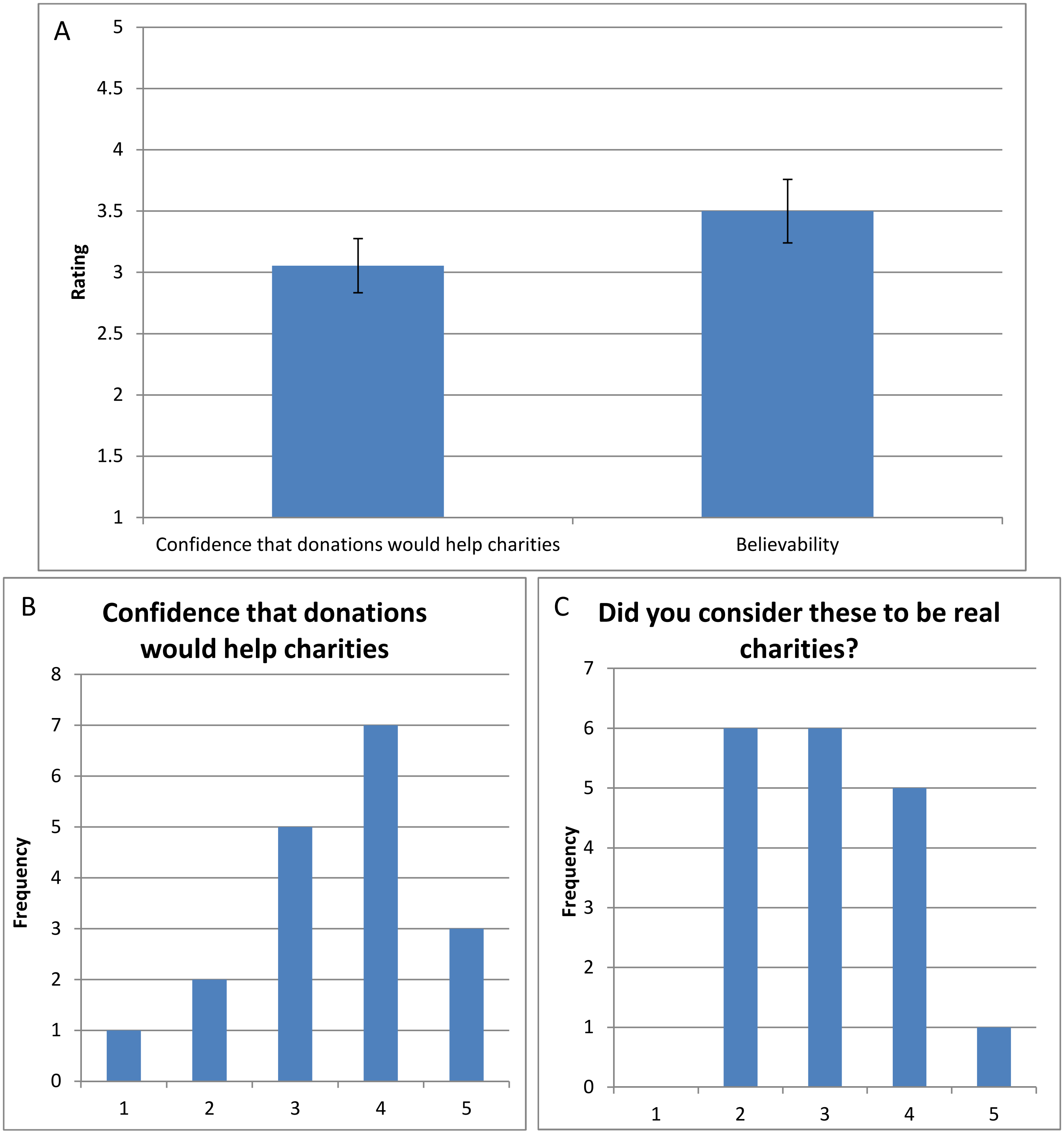

Figure 3. (A) Post scan ratings of the mean participant responses reflecting the believability of the task premise and the perceived helpfulness of donations made by the participant to the charities. (B)+(C) Frequency plots of the number of participants for each

Table 3. Areas demonstrating greater BOLD activation during Donation versus No Donation choices during the interstimulus interval

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Anticipation						
Help > Harm p < 0.005						
Middle Frontal Gyrus	L	8	-40.9	11.8	45.8	972
Middle Temporal Gyrus	L	21	-50.0	-32.3	-6.0	648
Superior Temporal Gyrus	L	38	-56.1	17.5	-10.3	270
Inferior Frontal Gyrus	L	47	-47.0	29.7	-6.0	270
Superior Frontal Gyrus	L	6	-4.5	26.6	59.6	270

3.3.2 Determining neural regions supporting decisions of restitution motivated by guilt

To determine whether acts of restitution reduce activation in regions implicated in the experience of guilt, we evaluated whole-brain BOLD signal during restitution responses. That is, when participants first did not initially donate to the individual in choice 1, but in the subsequent opportunity, chose to donate. We first used t-test to compare BOLD responses during a choice 2 donate decision that followed a no-donate decision for the same charity (Harm1/Help2) with BOLD responses during a choice 2 no-donate decision that followed a no-donate decision for the same charity (Harm1/Harm2) responses. This demonstrated increased BOLD signal during the restitutive action in the anterior cingulate, while greater activity for a non-restitutive decision (i.e. a second harm) was observed in middle and superior temporal gyri (Table 4).

To determine whether individual differences in guilt-proneness were associated with differential neural responses during this form of social decision making and restitution, in a follow up analysis, we performed the same voxel-wise contrast including the covariate of Trait Guilt scores from the Guilt Inventory. This demonstrated a positive correlation between trait guilt scores and BOLD activation in dorsomedial PFC (BA 24/32) and vlPFC (11/47), consistent with our predictions.

Table 4. Neural regions active during the Harm1/Help2 vs Harm1/Harm2 contrast

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Restitution > Harm, p < 0.005						
Anterior Cingulate	R	32	19.7	35.1	11.1	108
Harm > Restitution, p < 0.005						
Middle Temporal Gyrus	L	21	-59.1	-35.5	-2.6	189
Superior Temporal Gyrus	L	22	-59.1	10.7	3.3	108
Superior Temporal Gyrus	L	42	-62.1	-33.3	17.3	108

Table 5. Neural regions showing positive correlations with guilt inventory trait scores during the active Harm/Help vs Harm/Harm Contrast

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Restitution > Harm [Guilt Inventory] p < 0.005						
Ventrolateral Prefrontal Cortex	R	11/47	31.8	45.0	-1.6	162
Dorsomedial Prefrontal Cortex	L	24/32	-7.6	38.8	-1.9	54
Harm > Restitution [Guilt Inventory] p < 0.005						
Lingual Gyrus	L	18	-1.5	-88.0	-5.7	216
Superior Parietal Lobule	L	7	-10.6	-75.7	60.8	135

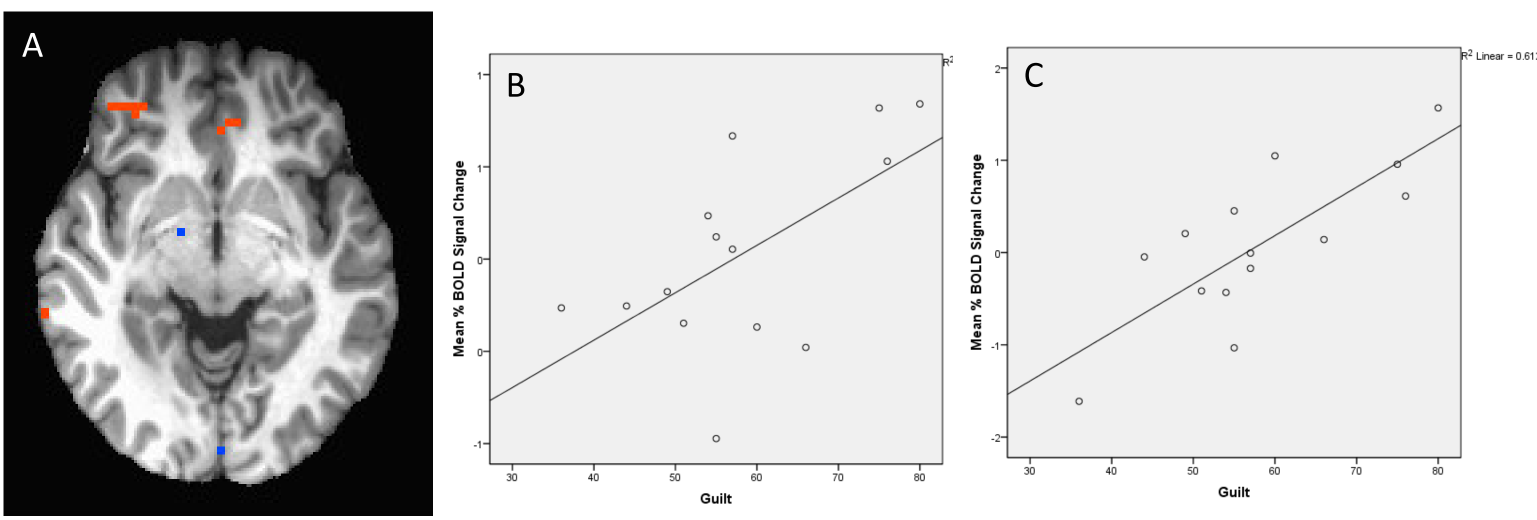

Figure 4. Increased vlPFC and dmPFC activation in guilt-prone individuals during acts of restitution. (A) Positive correlations are observed between the guilt inventory and vlPFC and dmPFC during acts of restitution. (B) Scatter plot of Trait Guilt scores from the Guilt Inventory with vlPFC activation, $R = 0.782$ ($p < 0.01$) and (C) with dmPFC activation, $R = 0.593$ ($p < 0.05$)

3.3.3 Identify neural regions active during choices to harm or help others

To identify regions of the brain supporting decisions to help vs. harm another individual, we first used t-tests to contrast BOLD responses during donations (Help1) with baseline. This demonstrated increased activity in regions that have been implicated in emotion processing and decision making such as the anterior cingulate and prefrontal cortices (Figure 5A). Next we conducted t-tests on BOLD responses during no donation decisions (Harm1) compared to baseline. This demonstrated increased activation in a network of regions partially overlapping that identified above for the help contrast including the frontal cortices, but notably the insula and temporal poles showed greater activity only during the choices of no donation (Figure 5B). Finally, we performed a direct contrast of Help1 vs. Harm1 choice phase BOLD responses using a voxel-wise t-test. This demonstrated significantly greater BOLD signal during harm decisions (no-donate) in the bilateral insula, amygdala and middle temporal gyrus (Figure 6 and Table 6).

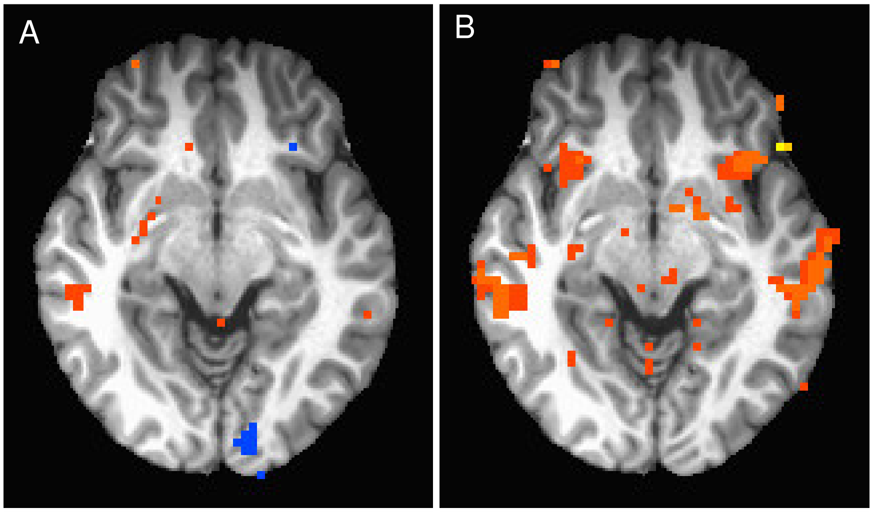

Figure 5. Neural regions showing significant BOLD responses compared to baseline during (A) donation choices (Help1), and (B) no-donation decisions (Harm1) suggesting that increased anterior insular activation may be specific to the harm condition

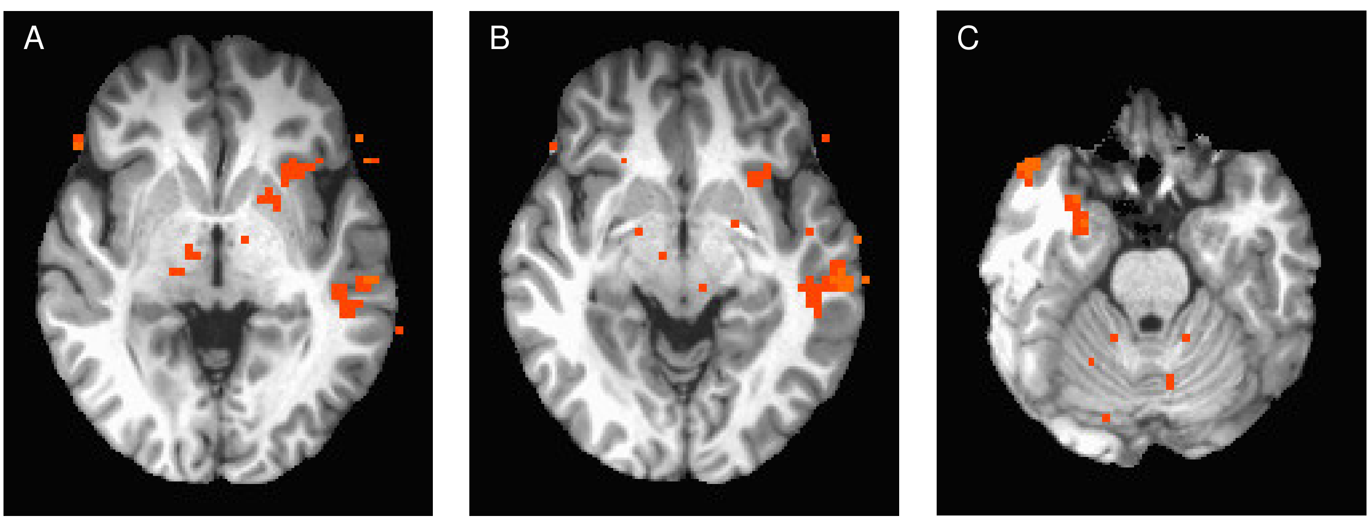

Figure 6. Brain regions demonstrating differential BOLD Responses during initial donate vs. no-donate decisions. Contrast between Harm vs Help. (A) Increased activity was observed in the (A) left anterior insula during no-donation decision (Harm1) when compared t to donations (Help2), (B) middle temporal gyrus, and (C) amygdala. Initial threshold contrast at $p<0.005$, corrected to $p<0.05$ for multiple comparisons

Table 6. Neural regions of interest in a contrast between Harm vs Help

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Harm > Help $p < 0.005$						
Middle Temporal Gyrus	L	21	-56.1	-23.0	-5.5	459
Insula	L		-25.8	17.2	-3.2	297
Amygdala	R		25.9	-0.4	-25.7	243

3.4 Active Condition versus Passive Condition Analysis

As active decisions that one makes which result in harm to another are associated with increased feelings of guilt compared to those that result in unintentional harm (Berthoz *et al.*, 2002 & Finger *et al.*, 2006), we examined the effect of agency upon BOLD signal activation. The contrasts between active and passive trials were intended to control for activity in response to the stimuli (pictures) and text. Using whole brain voxel-wise t-tests, we first compared BOLD signal during active choice phase when the participant made decisions to harm another individual (participant's choice) vs. the passive choices made by the computer to harm the individual (other's choice, participant simply pressed button to advance the screen) (Table 7). Of particular interest, the right amygdala, medial prefrontal cortex, and left insula showed greater activation during the active decision relative to the passive viewing of harm decisions being made (Figure 7). We then compared the active vs. passive choice phase for decisions to help another individual (Table 8). Of particular interest, the medial prefrontal cortex showed greater activation during the active phase when compared to the passive phase of helping acts (Figure 8).

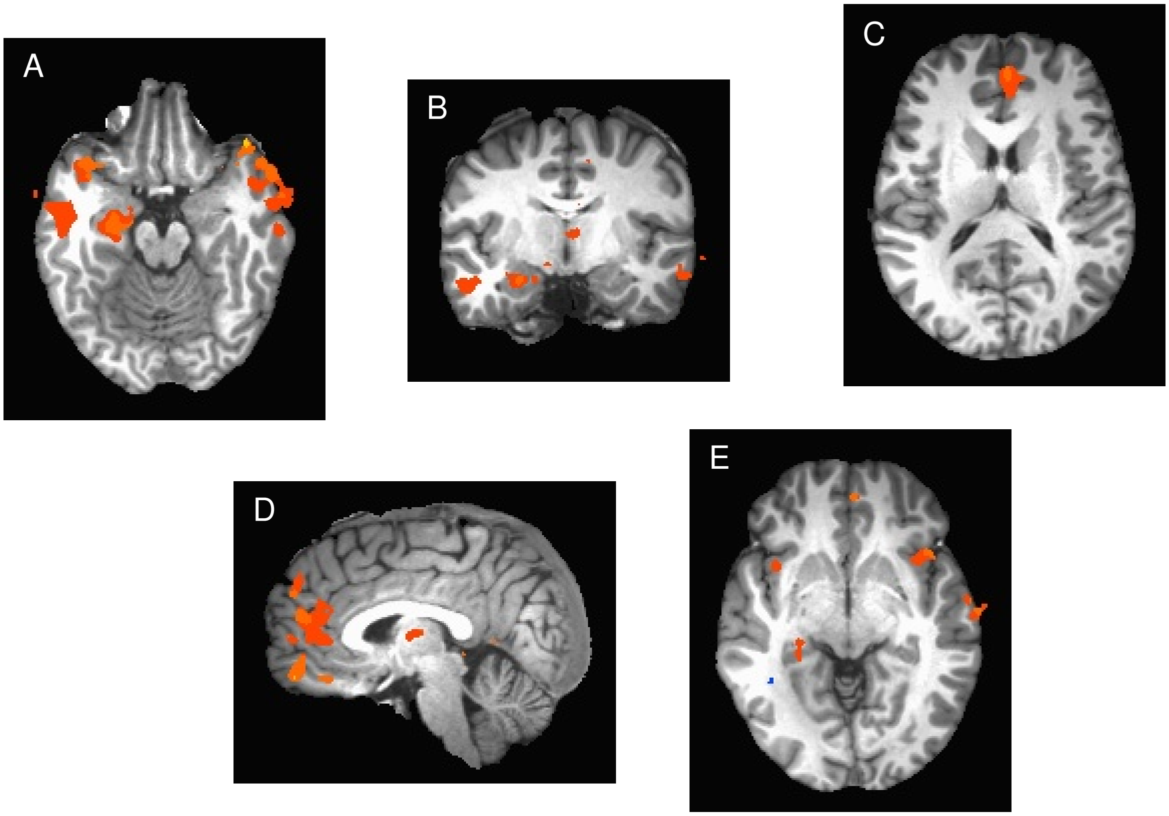

Figure 7. Regions demonstrating differential BOLD signal responses when comparing active versus passive conditions of choices to harm another individual. Greater BOLD responses were observed during active harm choices in the amygdala (A, B) and temporal poles (A), medial prefrontal cortex (C,D), and left anterior insula (E).

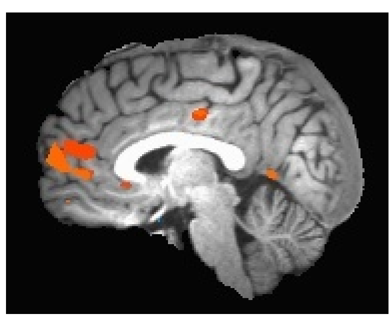

Figure 8. Increased ventromedial prefrontal cortex activation is observed during active relative to passive choice condition to help another individual

Table 7. Significantly active clusters from the voxel-wise t-test analysis of the active vs passive condition of decisions to harm another individual

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Active > Passive Harm1						
Superior Temporal Gyrus	L	38	-53.0	14.5	14.1	3996
Superior Frontal Gyrus	L	8	-13.6	48.7	51.0	3429
Middle Temporal Gyrus	R	21	50.0	-0.6	-22.1	2592
Medial Frontal Gyrus	L	32	-1.5	41.0	17.9	2052
Amygdala	R		25.8	-13.1	-19.2	1161
Insula	L	13/47	-37.9	20.3	-3.0	837
Medial Frontal Gyrus	L	10/11	-1.5	51.5	-8.3	621
Parahippocampal Gyrus	R	27	25.8	-26.1	-5.7	432
Thalamus	L		-1.5	-8.2	8.8	378
Cingulate Gyrus	R	32	10.6	18.4	36.3	378
Medial Frontal Gyrus	L	11	-4.5	36.2	-12.8	351
Cingulate Gyrus	L	32	-4.5	15.3	36.2	351
Lentiform	L		-13.6	-1.2	-7.8	270
Caudate	R		7.6	-2.6	22.2	243
Middle Temporal Gyrus	L	21	-62.1	-13.7	-5.0	216
Medial Frontal Gyrus	R	10	7.6	56.9	9.0	189
Caudate	L		-7.6	3.6	22.5	189
Insula	R	13	37.9	14.1	-3.4	162

Table 8. Significantly active clusters from the voxel-wise t-test analysis of the active vs passive condition of decisions to help another individual

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Active > Passive Help1						
Anterior Cingulate	L	32	-1.5	41.0	17.9	2808
Superior Temporal Gyrus	L	38	-56.1	8.2	-10.8	702
Superior Temporal Gyrus	R	38	59.1	11.6	-17.8	459
Middle Temporal Gyrus	R	21	50.0	-3.7	-22.3	270
Medial Frontal Gyrus	R	10	1.5	51.5	-8.3	270
Cingulate Gyrus	L	23	-1.5	-18.7	34.4	243
Superior Temporal Gyrus	L	38	-50.0	20.9	-17.3	216
Posterior Cingulate	L		-1.5	-54.3	3.1	189
Passive vs Active Help 1						
Inferior Parietal Lobe	L	40	-62.1	-43.5	36.3	567
Inferior Parietal Lobe	R	40	59.1	-34.3	36.8	162

3.5 Amplitude Modulated Analysis

Finally, in order to determine whether the BOLD signal in the functional data correlated significantly with subjective feelings of guilt, an amplitude modulated analysis was performed. In the analysis, all choices (harm or help) are collapsed and two main contrasts were performed. The first was a contrast between subjective trial-by-trial ratings of feelings of guilt after choice 1 versus baseline (zero), where baseline is defined in AFNI as the mean BOLD signal averaged across the entire duration of the run, to determine areas that are active to confirm the validity of the analysis. This contrast identified that the posterior cingulate, bilateral inferior parietal lobule, and the mPFC (see comment) (Table 9). The second contrast was to determine the difference between the effects of subjective feelings of guilt in the second choice phase compared to the first on BOLD signal. Here we hypothesized that dmPFC and vlPFC regions showing positive correlations with trait guilt ratings during restitution may also show positive correlations with this measure of state guilt. However, no significant amplitude modulation of BOLD signal by subjective trial-by-trial post-hoc ratings of guilt was found for these regions of PFC (Table 10).

Table 9. Neural regions with BOLD responses that are significantly modulated by feelings of guilt during choice 1 (harm or help) when compared to baseline

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Choice 1 > Baseline ratings; $p < 0.005$						
Precentral Gyrus	L	4	-31.8	-25.2	63.4	2646
Postcentral Gyrus	R	40	56.1	-31.9	53.3	1161
Middle Frontal Gyrus	L	6	-22.7	-17.0	63.9	567
Superior Frontal Gyrus	R	6	13.6	4.5	68.3	540
Lentiform Nucleus	L		-19.7	-1.5	-0.7	513
Cingulate Gyrus	R	31	10.6	-40.6	39.8	432
Medial Frontal Gyrus	L	6	-7.6	-10.	67.4	405
Precuneus	R	7	10.6	-71.3	34.9	378
Superior Temporal Gyrus	R	22	53.0	-39.0	7.2	270
Middle Frontal Gyrus	R	8/9	31.8	21.2	43.0	270
Inferior Parietal Lobe	R	40	40.9	-40.9	46.3	243
Inferior Parietal Lobe	L	7/40	-34.8	-50.6	55.5	216
Superior Temporal Gyrus	L	22	-56.1	-36.0	7.4	189
Post Central Gyrus	L	2/40	56.1	-28.9	53.4	189

Table 10. Neural regions that are significantly modulated by feelings of guilt during choice 2 when compared to the first choice

Structure	L/R	BA	X	Y	Z	Volume (mm^3)
Choice 2 > Choice 1 Guilt Ratings; $p < 0.005$						
Lentiform Nucleus	L		-19.7	-4.5	-4.4	351
Precentral Gyrus	R	6	16.7	-20.2	66.9	297
Precentral Gyrus	L	4	-34.8	-23.0	60.3	270
Middle Frontal Gyrus	L		-37.9	15.8	26.4	162
Choice 1 > Choice 2 Guilt Ratings						
Superior Frontal Gyrus	L	6/8	-22.7	11.5	52.3	189
Inferior Temporal Gyrus	L	20	-65.2	-22.4	-19.8	162

Chapter 4 Discussion

In the current study, we employed a novel social decision-making paradigm using fMRI to study the neural correlates of guilt and the subsequent restitutive behaviours. Based on rating scales and post-scan interviews, we found that participants were convinced of the believability of the study and subjectively experienced feelings of guilt. We observed increased BOLD signal during choices of harm when compared to help in the insula, STS, and amygdala suggesting potential functions consistent with proposed roles for these regions in processing emotional information and aversive cues, and anticipation of negative feedback. We also found that BOLD signal within the dmPFC and vlPFC was positively correlated with trait measures of guilt proneness during choices of restitution, supporting our hypothesis that these areas may play a role in altering behaviour in response to negative social feedback.

4.1 Behavioural results

Based on the post-hoc trial-by-trial subjective emotion ratings for guilt and compassion, the task successfully elicited these moral emotions, and in a pattern that is dissociable. As expected, when individuals did not donate, self-reported feelings of guilt increased. Feelings of compassion, on the other hand, showed the opposite trend, decreasing after choices of harm and increasing after choices of help. Although we did not make any specific predictions regarding the compassion ratings, this result is consistent with prior studies showing that individuals tend to help those who they believe are in need (Moll et al 2006). In line with past behavioural research (Ketelaar & Au, 2003), analysis of subjective trial by trial ratings of guilt demonstrated a decrease in guilt following acts of restitution. In contrast, initial feelings of guilt did not decrease following a second choice not to donate to the same individual. Additionally, although ratings for other emotions were not obtained, the decision patterns for trials that elicited guilt, a prosocial emotion that is experientially negative, were distinct from those of compassion, a prosocial emotion that is experientially positive (Tangney *et al.*, 1996).

4.2 Anticipatory Guilt

Contrary to our hypothesis and past research, we did not observe anterior insula activation in the contrast between the interstimulus intervals following no donations (harm) vs. donations (help). However, in the contrast comparing decisions of harm to help, during the first choice, we found increased bold signal within the anterior insula, as well as in the STS and amygdala. The finding of this activity pattern during the choice, rather than the subsequent interstimulus interval may indicate that the activity is related to prediction calculations at the time of decision making. Given the finding of increased anterior insula activity during harm relative to help choices, it appears that the insula does play a role in anticipating or calculating the potential negative consequences of a decision, at a time proximal to the decision.

The insula has been hypothesized to play a more general role in decision making, particularly in avoiding future negative feedback. There is also evidence showing that the insula may play a role in interoception and of emotional states. For example, Shin *et al.* (2000) has found that the anterior insula is activated during the re-experiencing of guilt inducing events. Additionally, Chang *et al.* (2011) found that activity in the insula was correlated with minimizing anticipated guilt in the future during the decision making phase of a trust game. However, in a prior related study by our group where participants made decisions to assign gains or losses to themselves or charities, we found that the anterior insula showed greatest activity when unintended outcomes, whether positive or negative, occurred (Greening *et al.*, 2013). Additionally, in that study, increased activity of the anterior insula during decision phase was correlated with fewer decisions to harm charities. We hypothesized, as a result, that during this form of social decision making, the anterior insula was more generally involved with risk-cost calculation rather than guilt specifically. In line with this, Bossaerts (2010) suggests that the anterior insula plays a role in risk tracking, where risk is the uncertainty that cannot be eliminated even after learning, and that such increased activation in anterior insula represents risk prediction error, that is the probability of making an error. He suggests that in response to risk predictions, the body "prepares" itself to respond to these uncertainties, which then, according to the James-Lange Somatic Marker Hypothesis, result in emotional states. The hypothesis states that physiological changes occur within the body, preceding the conscious awareness of emotions, and subsequently the mind then labels the emotion based upon the physiological changes and situational context. In support of this, in a study using the Cambridge Gamble Task, an insular cortex lesion group showed increased wagers during the task when compared to healthy participants, but in a pattern distinct from a vmPFC lesion group. The vmPFC lesion group showed increased wagers as well, but did so in accordance to the probability of winning; however, insular cortex lesion patients showed increased betting regardless of the probability of winning. This latter group was unable to track less favourable odds and performed the worst out of all groups, suggesting that the insular cortex is necessary for risk tracking and adjustment (Clark *et al.*, 2008).

In the current task we found increased amygdala activity during choices to not donate to charities. The classical view of the amygdala is that it is the fear centre of the brain, where its main role is in the acquisition and storage of fearful memories (LeDoux, 2000). It has been shown through aversive classical conditioning using acoustic stimuli in rats that lesions to the lateral amygdaloid nucleus of the amygdala prevent the conditioning and interfere with both the behavioural and autonomic emotional responses, effectively blocking fearful memory formation (LeDoux *et al.*, 1990). In humans suffering from bilateral amygdala lesions, it has been shown that there is a disability to recognize fearful facial expressions, without affecting recognition of facial identity (Adolphs *et al.*, 1994). This impairment is due to the failure to fixate gaze upon the eyes spontaneously, the most important feature to recognizing fear. This failure to recognize fear can be recovered when patients are instructed to fixate upon the eyes (Adolphs *et al.*, 2005). As the IAPS photos used include several stimuli with faces and facial expressions, it is possible that the increased activity observed in the amygdala in the present task may simply reflect the emotional stimuli of the IAPS. However, the amygdala is an area that has been implicated in a more general role in modulating the processing of emotional stimuli (Schwartz *et al.*, 2002; Cardinal *et al.,* 2002; Sergerie *et al.,* 2007; Heimer *et al.,* 2008). In a texture discrimination task, designed to test visual learning, when participants were looking at new pictures, they showed increased amygdala activity (Schwartz *et al.*, 2002). It is suggested that this increased functional connectivity may enhance visual learning through motivational signals, processing of relevant stimuli, and consolidation processes (Holland & Gallagher, 1999 & Vuilleumier *et al.*, 2001). Therefore it appears that the amygdala is not only activated during fearful learning or in response to fear, but also to detect emotionally salient but unpredictable events, regardless of attentional focus or load (Vuilleumier *et al.*, 2001), ultimately leading to shifting of attention to these stimuli and modulating behavioural responses and learning (Holland & Gallagher, 1999). In the current study, amygdala activity increased when contrasting between decisions of help and decisions of harm, which potentially may have been associated with different pictures. In the decisions to help, the pictures may have been more negative and therefore more salient and emotionally evocative for the participants, resulting in an increased likelihood to donate. Thus while it is possible that the IAPS stimuli elicited the amygdalar response, it is not likely the only factor driving this response, given that increased activity was seen during choices of harm and not help.

Therefore the amygdala activity observed may not simply reflect response to the IAPS, but rather seems to be tied to the decision to not donate as well.

In line with this literature, deficits in empathic perspective taking and emotion attribution are correlated with atrophy in the amygdala, along with the insula, STS and regions of the temporal poles and frontal lobes in frontotemporal dementia (Eslinger *et al.*, 2011 & Cerami *et al.*, 2014). This suggests that these areas within the frontoinsular network play an important role in supporting empathy, which is considered a critical emotion for prosocial decision-making. Alternatively, a more general role for the amygdala in reflecting the salience of a stimulus or situation has been proposed. Bossaerts (2010) suggests that the amygdala plays a role in tracking estimation uncertainty, which represents the uncertainty that can be reduced with additionally learning. Adams et al. (2003) also showed that ambiguity resulted in greater amygdala activation, when studying angry and fearful facial expressions by manipulating gaze in the facial expressions. Fearful expressions with direct gaze and angry expressions with averted gaze elicited greater activity in the amygdala, in both of these, the source of threat is unknown to the participant and is thus more produces ambiguity. Similarly, this amygdalar activation has been replicated in potentially ambiguous facial expressions of fear, sadness, happiness, and anger, thus further outlining the role of the amygdala not only in fear but more general emotional processing and ambiguity disentangling (Yang *et al.*, 2002). Thus the amygdala has a potential role in determining biological and social relevant information (Sander *et al.*, 2003 & Sergerie *et al.*, 2007).

Prior studies looking at guilt and embarrassment have found activation within the left STS for both of these social emotions (Takahashi *et al.*, 2004 & Finger *et al.*, 2006). The STS is typically activated across a variety of tasks featuring social stimuli or social interactions

(Takahashi *et al.*, 2004, Singer *et al.*, 2004, Harris *et al.*, 2005, Finger *et al.*, 2006, Lee & Harris, 2013). It has been proposed that the STS processes social cues that reveal communicative intentions of others through the analysis of eye-gaze direction, facial expressions, body movements, and other types of biological motions (Redcay, 2008). In our current study, we hypothesize that the STS activity reflects processing of the facial expression and ultimately the intention of the subject within the IAPS stimuli. In conjunction with the amygdala, we suggest this information from the STS regarding the goals and intentions of the individuals and how they are in need contributes to ToM processing and empathy for the depicted individuals, and is integrated into cost/benefit calculations/predictions in the insula and amygdala.

4.3 Neural correlates of restitution: Roles of vlPFC and dmPFC

When restitution was contrasted with choices of harm, that is comparing Harm followed by Help choices (Harm/Help) with Harm followed by Harm choices (Harm/Harm), we found increased activation in the anterior cingulate cortex (ACC, BA 32).The ACC is an area that, in the past, has been referenced to as the dmPFC, although the ACC can include broader and more ventral prefrontal regions to the dmPFC (Venkatraman & Huettel, 2013). The ACC and neighboring regions of mPFC have long been implicated in decision-making, emotional awareness, and risk uncertainty and anticipation (Craig, 2009). Of interest, the ACC is typically co-activated with the insula during interoception, particularly during viewing of fearful faces, and perceptual decision making (Craig, 2009). This division of the medial frontal cortex has been suggested to use reward and other outcome information to guide future behaviours (Walton *et al.*, 2004 & Rushworth et al., 2005). Waytz et al. (2011) also found that the dmPFC predicted altruistic behaviours, where highest activity was correlated with the greatest amount of money donated to others.

To test the hypothesis that persons more prone to guilt may experience greater drives towards restitutive actions, we included a measure of Trait Guilt as a covariate in the restitution contrast. We found a positive correlation with trait guilt scores and activity specifically in vlPFC and mPFC (anterior cingulate cortex) during the acts of restitution. That is, in participants that are more prone to guilt, there is greater activation in these two areas. These results are consistent with and extend results from a previous study in the lab that used a guilt-inducing passive sentence viewing task and found that these two areas became activated during moral and social transgressions, which also was correlated with feelings of guilt. It was hypothesized that these areas are important for processing aversive social information in order to facilitate a behavioural change from the one that elicited the negative feedback to a more advantageous option (Finger et al., 2006). More recently, Morey *et al.* (2012) employed a similar task where participants read hypothetical scenarios that were designed to induce guilt as a result of harming another or harming yourself. Participants then rated their feelings of guilt subsequent to reading the stories, and they found that guilt intensity was correlated with the dmPFC, vlPFC, frontopolar cortex, and supramarginal gyrus.

These roles for dmPFC and vlPFC in social decision making and response selection are supported by the literature regarding their role in across a variety of social and non-social cognitive tasks. The dmPFC has been found show activity during tasks designed to induce erroneous responses, such as the counting Stroop (Bush et al., 1998), Continuous Performance Test (Carter et al., 1998), and the flanker task (Botvinick et al., 1999), and greater activity is observed with increasing degrees of conflict between the decisions. This research has suggested a role of the dmPFC in monitoring errors. Subsequent studies have shown that this signal triggers behavioural or strategic changes that prevent subsequent errors (Ullsperger and von Cramon, 2003; Brown and Braver, 2005; Kennerly *et al.*, 2006). The dmPFC has also been found to be activated during a charity task, when assigning a loss to the charity relative to assigning a loss to oneself. It was suggested this increased dmPFC activity may reflect greater decision conflict for decisions to harm another (Greening *et al.*, 2013). Similar to the dmPFC, the vlPFC has been found to be activated in response to negative feedback that prompts behavioural changes such as in reversal learning paradigms (Cool *et al.*, 2002; Mitchell *et al.* 2009).

Thus, results from the present study support our previous finding of increased activation in dmPFC and vlPFC during guilt-inducing scenarios and extend the hypothesis to suggest that, particularly in guilt prone individuals, feedback depicting harm to another is processed similar to an error in dmPFC, which in turn primes vlPFC to modify the causative behaviour. Of interest is why individuals with greater guilt proneness may therefore show greater activation in these areas- i.e. how is a greater "social error signal" generated in dmPFC in some individuals. We propose that guilt prone individuals are more responsive to aversive social feedback and are more likely to engage in subsequent behavioural changes as a result of increased empathy or differences in risk prediction calculations, which may relate to potential feelings of guilt. However, we did not observe significant correlations between the guilt inventory and the neural regions such as STS, amygdala and insula implicated in processing the emotional state and needs of the charity individuals. This lack of signal may be secondary to insufficient power, as the participants were a community sample, not specifically recruited for high or low guilt trait scores. An alternate model, though we suggest a less likely one, may be that as opposed to the individual's guilt proneness resulting in greater error signals, it could be that the greater error signals prompt increased feelings of empathy. Further investigation to confirm which model may be more accurate could be addressed by recruiting two more extreme populations, one that is high in guilt proneness and one that is low, and contrasting the BOLD signal during feedback following a harm decision and during restitutive decisions. This would allow us to further determine whether increased activity is observed in the amygdala, STS, and insula for individuals high in guilt-proneness, and whether these signals correlate with subsequent signal in the vlPFC and dmPFC in the subsequent decision phase.

4.4 Amplitude Modulated Analysis

Subjective ratings of guilt showed that participants felt significantly less guilty after restitutive behaviours. The goal of the amplitude modulated analysis was to determine whether subjective feelings of guilt would modulate the BOLD activity within any regions of interest, on a trial by trial basis However, using a whole-brain voxel-wise approach we did not find any significant correlation with the subjective guilt ratings and BOLD signal in our a-priori areas of interest including the insula, amygdala, and STS. The lack of results may have been due to potential unreliability of the post-hoc explicit recall of subconscious feelings on over 100 trials during the scanning procedure. This problem could be addressed in future studies by including more objective autonomic measures of emotional states in real-time such as heart rate or skin conductance. These measures have been used in paradigms that have been designed to study guilt and have found physiologic arousal when experiencing guilt (Eisenberg *et al.*, 1991; Gamer *et al.*, 2008; Wahlund *et al.*, 2010; Oliveira-Silva & Goncalves 2011; Kouchaki *et al.*, 2013). Although these measures may not directly be able to identify guilt specifically and are related more to general arousal, it is possible to use the differences between these measures and the participants' baseline to determine neural patterns that may reflect internal emotional states.

4.5 Limitations, potential pitfalls and confounds

Guilt is a complex social emotion that is often grouped with two other emotions, shame and embarrassment, due to various similarities in the situations upon which they are elicited. As a result, there is always the potential confound that the emotion that the paradigm was designed to isolate may result in overlap with these other social emotions. A potential way to disentangle these emotions may be to specify whether others have knowledge of their actions, as guilt should occur in response to harm to a victim, whether or not it is witnessed, whereas shame may vary as a function of others knowledge of one's action. Another point that may be of concern is the success of the deception used in our task and the difficulty in replicating "real life" situations and emotions while within the MRI scanner. Although participants' post scan interview ratings showed that some participants, while not necessarily believing the story of the individuals in need, still believed in the existence of the charities and that donations would actually help those in need. However, past studies have found that individuals who role play often "take on" the emotions and identities of their role (Haney *et al.*, 1973). Thus although there are participants who fall below the median value of the believability scale, it does not necessarily mean they were not generating the emotions or evoking the same neural regions.

Another complicating factor in the design of the study linked decisions of harm towards another individual with benefits towards the self. Thus during the decision making process, factors aside from guilt or empathy, such as considering one's own needs and self-interest could potentially influence and confound the final decisions. However, this design was selected for its ecological validity as decisions to help another typically involve a cost to oneself. A final point that may be of concern is the power of the results, specifically, in the event number for the trials of restitution, which was lower than the other possible outcomes of choices. The current study used a novel task design and the current behavioural results show that participants engaged in the task experienced guilt and engaged in restitutive behaviours in response to guilt. The results of the study, however, show an imbalance in number of responses in the four possible categories of behaviours (Help-Help, Help-Harm, Harm-Harm, Harm-Help), which may affect the power of the statistical analyses. Specifically, there was a significantly lesser amount of responses for the restitutive behaviours (Harm/Help: 9.59 +/- 2.24 vs Help/Help: 13.76 +/- 2.85, Help/Harm: 26.41 +/- 2.51, Harm/Harm: 18.94 +/- 1.86). Thus in order to increase these responses, providing stronger negative feedback in response to decisions to not donate to charities may be a potential way to balance the responses. As we did not see significant correlations between the post-hoc trial-by-trial guilt ratings and BOLD signal in regions such as those active during harm decisions, it may be of consideration to ask participants to rate feelings of guilt subsequent to each story, immediately after making the donation decision, in order to reduce confounds related to post-hoc memory of a emotions for a large number of trials. This would also reduce the potential for to manipulate answers in hindsight, when shown their original donate or no-donate responses.

Another limitation is the predictability of the task, where the participants may have learned that they would always be presented with two opportunities to donate. As a result, participants may have made decisions for both choice 1 and choice 2 prior to the actual choice phase, which could obscure specific neural signals related to guilt and restitution. A potential way to resolve this issue in future studies would be to alter the design to be less predictable by only offering a second choice randomly and by varying the dollar amount for donations in both the choices.

Although only the Guilt Inventory showed correlation with the imaging data, this result and possibly the lack of correlation with the other trait measures may be a result of gender-specific differences. Although there are no studies that have studied-specific differences in the experience of guilt or compassion, there are different sets of normalized results for males and females within these trait scales, it is possible that the intensity of ratings may differ, which may then also be reflected in activation intensity. However, the current study was not designed to look at gender-specific activations, therefore it is not powered to do so, but would of interested for future studies.

4.6 Future Directions

A potential follow up analysis for the current study would be functional connectivity analysis, which determines whether the areas of interest show are functionally connected by analyzing the time course of activation during the task. Additionally, another follow up analysis would be the use of multivariate pattern analysis (MVPA) and a support vector machine (SVM). SVMs are used to predict behavioural choices based upon a subset of the data, while the MVPA provides patterns of activation that can be used to entrain the SVM. In our current study, this analysis would serve to further allow us to conclude that the areas of the brain that were found to be correlated with resitutive behaviours support these behaviours. That is, based on the patterns of activation observed during choice and feedback one, if the SVM is able to predict the decisions of the participants for choice 2 correctly, then a conclusion towards the BOLD signal observed in these areas of interest supporting these behaviours is much stronger.

We found a correlation with guilt and restitution related activation in dmPFC and vlPFC, but no significant restitution-related activity in these regions in the absence of the covariate. This suggests that power to detect such activity may be increased by examining healthy controls that score on the extreme ends of the spectrum. This would also allow us to determine whether the guilt inventory result replicates in the population that scores higher on the scale.

One of the goals of the study was to determine the neural correlates of guilt and restitution, both of which are lost early on in the disease progression of FTD and psychopathic personality disorder. It will be of interest to replicate the study in individuals suffering from psychopathic personality disorder to observe the changes in the BOLD signal responses or atrophy in the areas of interest (amygdala, insula, STS) identified in harm decision in the current study. It is expected that there will be a decrease in activation within these regions, due to the lack of empathy within this population. Given their empathy deficits, it will be unlikely that these individuals will engage in restitutive behaviours, thus the vlPFC and dmPFC results may not be replicable within this population.

4.7 Conclusions

The present study describes a novel fMRI paradigm that provided participants the opportunity to actively make decisions that would help or harm a person in need, and which featured feedback designed to elicit feelings of guilt. The amygdala, insula, and STS were found to be activated during choices of harm when compared to choices of help. We also demonstrated that the vlPFC and dmPFC may be key regions supporting decisions of restitution. Our findings support and extend prior models of social decision-making, where it is proposed that the STS and amygdala work in conjunction to support empathy through their functions in processing emotional expressions, other's intentions and modulating emotional salience, feeding this information to the anterior insula to predict future outcomes of behaviours. When a decision resulting in negative feedback is made, if weighted as sufficiently costly to the individual, the vlPFC and dmPFC are engaged and facilitate a change in response to alleviate the harm done when possible.

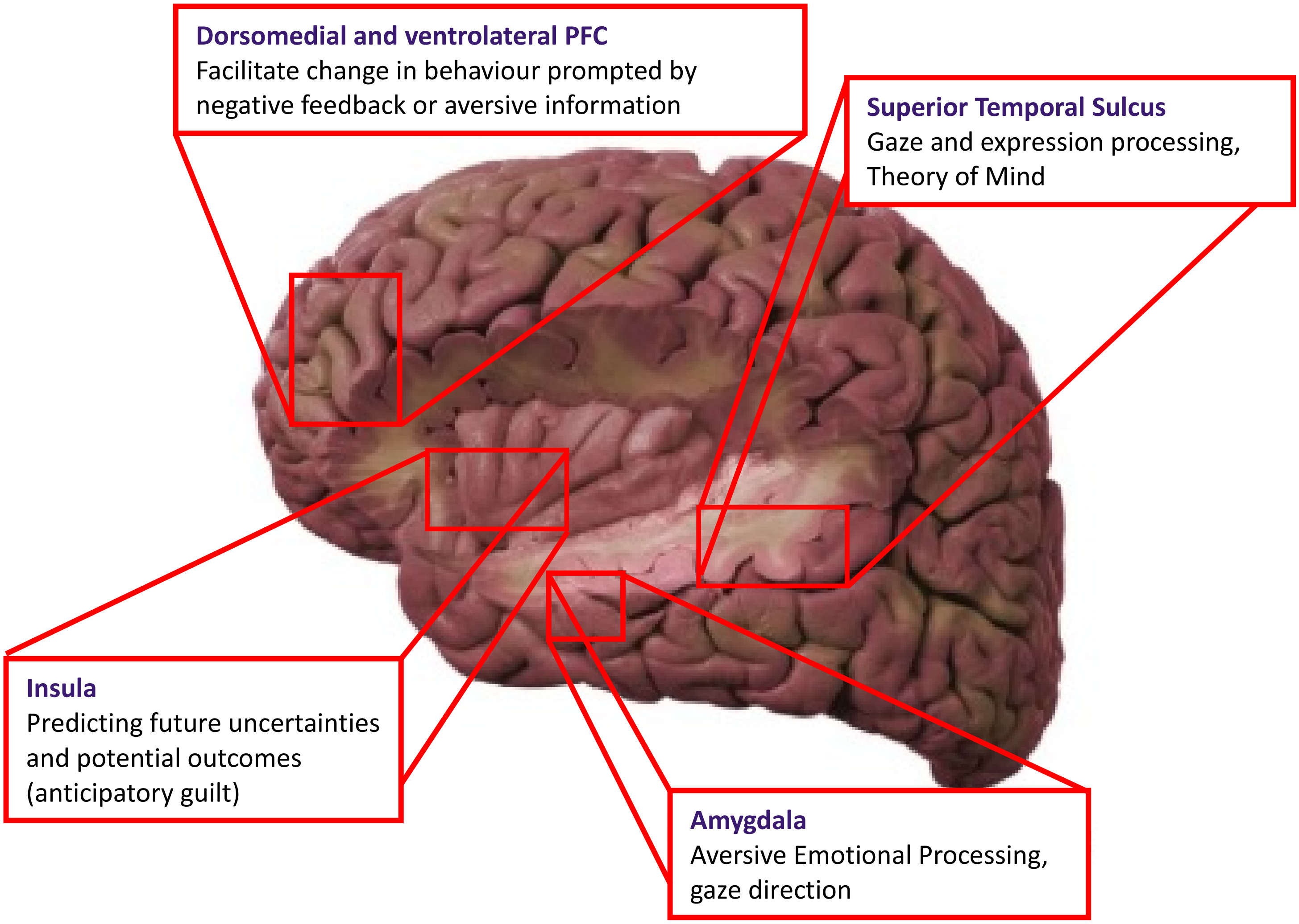

Figure 7. Hypothesized model of social decision making and restitution

References

Adolphs, R. (1999). The Human Amygdala and Emotion. *The Neuroscientist*, *5*(2), 125–137. doi:10.1177/107385849900500216

Adolphs, R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 60–5. doi:10.1038/nature03051

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716. doi:10.1146/annurev.psych.60.110707.163514

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 68–72. doi:10.1038/nature03086

Adolphs, R., Russell, J. a., & Tranel, D. (1999). A Role for the Human Amygdala in Recognizing Emotional Arousal From Unpleasant Stimuli. *Psychological Science*, *10*(2), 167–171. doi:10.1111/1467-9280.00126

Adolphs, R., Tranel, D., Damasio, H., & Damasio, a R. (1995). Fear and the human amygdala. *The Journal of Neuroscience□: The Official Journal of the Society for Neuroscience*, *15*(9), 5879–91.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*, 669–672.

Adolphs, R., Tranel, D., Hamann, S., Young, a W., Calder, a J., Phelps, E. a, … Damasio, a R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia*, *37*(10), 1111–7.

Anderson, a K., & Phelps, E. a. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*(6835), 305–9. doi:10.1038/35077083

Anderson, S. W., Bechara, a, Damasio, H., Tranel, D., & Damasio, a R. (1999a). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, *2*(11), 1032–7. doi:10.1038/14833

Anderson, S. W., Bechara, a, Damasio, H., Tranel, D., & Damasio, a R. (1999b). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, *2*(11), 1032–7. doi:10.1038/14833

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: an interpersonal approach. *Psychological Bulletin*, *115*(2), 243–67.

Beckmann, M., Johansen-Berg, H., & Rushworth, M. F. S. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *The Journal of Neuroscience□: The Official Journal of the Society for Neuroscience*, *29*(4), 1175–90. doi:10.1523/JNEUROSCI.3328-08.2009

Berndsen, M., van der Pligt, J., Doosje, B., & Manstead, A. (2004). Guilt and regret: The determining role of interpersonal and intrapersonal harm. *Cognition & Emotion*, *18*(1), 55–70. doi:10.1080/02699930244000435

Berthoz, S., Armony, J. L., Blair, R. J. R., & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain□: A Journal of Neurology*, *125*(Pt 8), 1696–708.

Blair, R. J., & Cipolotti, L. (2000). Impaired social response reversal. A case of "acquired sociopathy". *Brain□: A Journal of Neurology*, *123 (Pt 6*, 1122–41.

Blair, R. J., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain□: A Journal of Neurology*, *122 (Pt 5*, 883–93.

Bossaerts, P. (2010). Risk and risk prediction error signals in anterior insula. *Brain Structure and Function*, 645–653. doi:10.1007/s00429-010-0253-1

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*(6758), 179–81. doi:10.1038/46035

Brown, J., & Braver, T. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118–1121.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: an interference task specialized for functional neuroimaging--validation study with functional MRI. *Human Brain Mapping*, *6*(4), 270–82.

Calder, a J., Keane, J., Manes, F., Antoun, N., & Young, a W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*(11), 1077–8. doi:10.1038/80586

Cardinal, R. N., Parkinson, J. a, Hall, J., & Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews*, *26*(3), 321–52.

Carlo, G., Mcginley, M., Davis, A., & Streit, C. (n.d.). Behaving badly or goodly: Is it because I feel guilty, shameful, or sympathetic? Or is it a matter of what I think?, (136), 75–93. doi:10.1002/yd

Carnì, S., Petrocchi, N., Del Miglio, C., Mancini, F., & Couyoumdjian, A. (2013). Intrapsychic and interpersonal guilt: a critical review of the recent literature. *Cognitive Processing*. doi:10.1007/s10339-013-0570-4

Carter, C. S., Braver, T., Barch, D. M., Botvinick, M., Noll, D., & Cohen, J. D. (1998). Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. *Science*, *280*(5364), 747–749. doi:10.1126/science.280.5364.747

Cerami, C., Dodich, A., Canessa, N., Crespi, C., Marcone, A., Cortese, F., … Cappa, S. F. (2014). Neural correlates of empathic impairment in the behavioral variant of frontotemporal dementia. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, 1–8. doi:10.1016/j.jalz.2014.01.005

Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, *70*(3), 560–72. doi:10.1016/j.neuron.2011.02.056

Chang, L. J., Yarkoni, T., Khaw, M. W., & Sanfey, A. G. (2013). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cerebral Cortex (New York, N.Y. : 1991)*, *23*(3), 739–49. doi:10.1093/cercor/bhs065

Christensen, J. F., & Gomila, a. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: a principled review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1249–64. doi:10.1016/j.neubiorev.2012.02.008

Clark, L., Bechara, a, Damasio, H., Aitken, M. R. F., Sahakian, B. J., & Robbins, T. W. (2008). Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain : A Journal of Neurology*, *131*(Pt 5), 1311–22. doi:10.1093/brain/awn066

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. (2005). Searching for the pan-cultural core of psychopathic personality disorder. *Personality and Individual Differences*, *39*(2), 283–295. doi:10.1016/j.paid.2005.01.004

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(11), 4563–7. doi:20026435

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

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *The Journal of Comparative Neurology*, *493*(1), 154–66. doi:10.1002/cne.20749

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

Cryder, C. E., Springer, S., & Morewedge, C. K. (2012). Guilty feelings, targeted actions. *Personality & Social Psychology Bulletin*, *38*(5), 607–18. doi:10.1177/0146167211435796

De Hooge, I. E., Nelissen, R. M. a, Breugelmans, S. M., & Zeelenberg, M. (2011). What is moral about guilt? Acting “prosocially” at the disadvantage of others. *Journal of Personality and Social Psychology*, *100*(3), 462–73. doi:10.1037/a0021459

Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Human Brain Mapping*, *25*(1), 22–34. doi:10.1002/hbm.20127

Eisenberg, N. (2000). E MOTION , R EGULATION , AND M ORAL, 665–697.

Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cerebral Cortex (New York, N.Y. : 1991)*, *10*(3), 308–17.

Eslinger, P. J., Moore, P., Antani, S., Anderson, C., & Grossman, M. (2012). Apathy in frontotemporal dementia: behavioral and neuroimaging correlates. *Behavioural Neurology*, *25*(2), 127–36. doi:10.3233/BEN-2011-0351

Eslinger, P., & Moore, P. (2011). Social cognition, executive functioning, and neuroimaging correlates of empathic deficits in frontotemporal dementia. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *23*(1), 74–81.

Fellows, L. K. (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Annals of the New York Academy of Sciences*, *1239*, 51–8. doi:10.1111/j.1749-6632.2011.06229.x

Finger, E. C., Marsh, A. a, Kamel, N., Mitchell, D. G. V, & Blair, J. R. (2006). Caught in the act: the impact of audience on the neural response to morally and socially inappropriate behavior. *NeuroImage*, *33*(1), 414–21. doi:10.1016/j.neuroimage.2006.06.011

Finger, E. C., Marsh, A. a, Mitchell, D. G., Reid, M. E., Sims, C., Budhani, S., … Blair, J. R. (2008). Abnormal ventromedial prefrontal cortex function in children with psychopathic traits during reversal learning. *Archives of General Psychiatry*, *65*(5), 586–94. doi:10.1001/archpsyc.65.5.586

FLYNN, F. G. (1999). Anatomy of the insula functional and clinical correlates. *Aphasiology*, *13*(1), 55–78. doi:10.1080/026870399402325

Frick, P. J., & White, S. F. (2008). Research review: the importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(4), 359–75. doi:10.1111/j.1469-7610.2007.01862.x

Gamer, M., Gödert, H. W., Keth, A., Rill, H.-G., & Vossel, G. (2008). Electrodermal and phasic heart rate responses in the Guilty Actions Test: comparing guilty examinees to informed and uninformed innocents. *International Journal of Psychophysiology□: Official Journal of the International Organization of Psychophysiology*, *69*(1), 61–8. doi:10.1016/j.ijpsycho.2008.03.001

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

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523. doi:10.1016/S1364-6613(02)02011-9

Greening, S., Norton, L., Virani, K., Ty, A., Mitchell, D., & Finger, E. (2013). Individual differences in the anterior insula are associated with the likelihood of financially helping versus harming others. *Cognitive, Affective & Behavioral Neuroscience*. doi:10.3758/s13415-013-0213-3

Gregory, S., Ffytche, D., Simmons, A., Kumari, V., Howard, M., Hodgins, S., & Blackwood, N. (2012). The antisocial brain: psychopathy matters. *Archives of General Psychiatry*, *69*(9), 962–72. doi:10.1001/archgenpsychiatry.2012.222

Haney, C., Banks, C., & Philip, Z. (1973). Interpersonal dynamics in a simulated prison. *International Journal of Criminology and Penology*, *1*, 69–97.

Harris, L. T., Todorov, A., & Fiske, S. T. (2005). Attributions on the brain: neuro-imaging dispositional inferences, beyond theory of mind. *NeuroImage*, *28*(4), 763–9. doi:10.1016/j.neuroimage.2005.05.021

Herpertz, S. C., Werth, U., Lukas, G., Qunaibi, M., Schuerkens, a, Kunert, H. J., … Sass, H. (2001). Emotion in criminal offenders with psychopathy and borderline personality disorder. *Archives of General Psychiatry*, *58*(8), 737–45.

Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes, *3*(2), 65–73.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–7. doi:10.1038/nn1724

Ketelaar, T., & Tung Au, W. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition & Emotion*, *17*(3), 429–453. doi:10.1080/02699930143000662

Kouchaki, M., Gino, F., & Jami, A. (2013). The Burden of Guilt: Heavy Backpacks, Light Snacks, and Enhanced Morality. *Journal of Experimental Psychology. General*. doi:10.1037/a0031769

Laible, D., Eye, J., & Carlo, G. (2008). Dimensions of Conscience in Mid-adolescence: Links with Social Behavior, Parenting, and Temperament. *Journal of Youth and Adolescence*, *37*(7), 875–887. doi:10.1007/s10964-008-9277-8

LeDoux, J. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Ledoux, J. E., Romanski, M., & Xagoraris, A. (1990). The Lateral Amygdaloid in Fear Conditioning Nucleus: Sensory Interface Amygdala, (April).

Lee, V. K., & Harris, L. T. (2013). How social cognition can inform social decision making. *Frontiers in Neuroscience*, *7*(December), 259. doi:10.3389/fnins.2013.00259

Lerner, M. J. Handbook of PSYCHOLOGY (Vol. 5). Hoboken: New Jersey (2003).

Lindsey, L. L. M. (2005). Anticipated Guilt as Behavioral Motivation, *31*(4), 453–481.

Lindsey, L. L. M., & Hill, J. B. (2007). Anticipated Guilt as Motivation to Help Unknown Others: An Examination of Empathy as a Moderator. *Communication Research*, *34*(4), 468–480. doi:10.1177/0093650207302789

Marsh, A. a. (2013). What can we learn about emotion by studying psychopathy? *Frontiers in Human Neuroscience*, *7*(May), 181. doi:10.3389/fnhum.2013.00181

Martino, B. De, Kumaran, D., Seymour, B., & Dolan, R. J. (2009). Europe PMC Funders Group Frames , Biases , and Rational Decision-Making in the Human Brain, *313*(5787), 684–687. doi:10.1126/science.1128356.Frames

Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An investigation of moral judgement in frontotemporal dementia. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, *18*(4), 193–7.

Mendez, M. F., McMurtray, A., Licht, E., Shapira, J. S., Saul, R. E., & Miller, B. L. (2006). The scale for emotional blunting in patients with frontotemporal dementia. *Neurocase*, *12*(4), 242–6. doi:10.1080/13554790600910375

Mendez, M. F., & Shapira, J. S. (2011). Loss of emotional insight in behavioral variant frontotemporal dementia or “frontal anosodiaphoria”. *Consciousness and Cognition*, *20*(4), 1690–6. doi:10.1016/j.concog.2011.09.005

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167

Mitchell, D. G. V. (2011). The nexus between decision making and emotion regulation: a review of convergent neurocognitive substrates. *Behavioural Brain Research*, *217*(1), 215–31. doi:10.1016/j.bbr.2010.10.030

Mitchell, D. G. V, Luo, Q., Avny, S. B., Kasprzycki, T., Gupta, K., Chen, G., … Blair, R. J. R. (2009). Adapting to dynamic stimulus-response values: differential contributions of inferior frontal, dorsomedial, and dorsolateral regions of prefrontal cortex to decision making. *The Journal of Neuroscience□: The Official Journal of the Society for Neuroscience*, *29*(35), 10827–34. doi:10.1523/JNEUROSCI.0963-09.2009

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

Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., & Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(42), 15623–8. doi:10.1073/pnas.0604475103

Moll, J., Zahn, R., & Oliveira-Souza, R. de. (2011). Impairment of prosocial sentiments is associated with frontopolar and septal damage in frontotemporal dementia. *Neuroimage*, *54*(2), 1735–1742. doi:10.1016/j.neuroimage.2010.08.026.Impairment

Morey, R. a, McCarthy, G., Selgrade, E. S., Seth, S., Nasser, J. D., & LaBar, K. S. (2012). Neural systems for guilt from actions affecting self versus others. *NeuroImage*, *60*(1), 683–92. doi:10.1016/j.neuroimage.2011.12.069

Nachev, P., Rees, G., Parton, A., & Kennard, C. (2005). Volition and conflict in human medial frontal cortex, *15*(2), 122–128. doi:10.1016/j.cub.2005.01.006.Volition

Niedenthal, P. (2007). Embodying emotion. *Science*, *316*(5827), 1002–1005.

Oliveira-Silva, P., & Gonçalves, O. F. (2011). Responding empathically: a question of heart, not a question of skin. *Applied Psychophysiology and Biofeedback*, *36*(3), 201–7. doi:10.1007/s10484-011-9161-2

Pastötter, B., Gleixner, S., Neuhauser, T., & Bäuml, K.-H. T. (2013). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition*, *126*(3), 373–7. doi:10.1016/j.cognition.2012.11.003

Paul, J. (2011). Social Cognition , Executive Functioning , and Neuroimaging Correlates ...

Pessoa, L. (2010). Emotion and cognition and the amygdala: from “what is it?” to “what’s to be done?”. *Neuropsychologia*, *48*(12), 3416–29. doi:10.1016/j.neuropsychologia.2010.06.038

Radua, J., Sarró, S., Vigo, T., Alonso-Lana, S., Bonnín, C. M., Ortiz-Gil, J., … Pomarol-Clotet, E. (2013). Common and specific brain responses to scenic emotional stimuli. *Brain Structure & Function*. doi:10.1007/s00429-013-0580-0

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neuroscience and Biobehavioral Reviews*, *32*(1), 123–42. doi:10.1016/j.neubiorev.2007.06.004

Rogstad, J. E., & Rogers, R. (2008). Gender differences in contributions of emotion to psychopathy and antisocial personality disorder. *Clinical Psychology Review*, *28*(8), 1472–84. doi:10.1016/j.cpr.2008.09.004

Rushworth, M. F. ., & Walton, M. E. (n.d.). Cognitive Neuroscience: Resolving Conflict in and over the Medial, 54–56.

Rushworth, M. F. S., Kennerley, S. W., & Walton, M. E. (2005). Cognitive neuroscience: resolving conflict in and over the medial frontal cortex. *Current Biology: CB*, *15*(2), R54–6. doi:10.1016/j.cub.2004.12.054

Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Reviews in the Neurosciences*, *14*(4), 303–16.

Sanfey, A. G., Rilling, J. K., Aronson, J. a, Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science (New York, N.Y.)*, *300*(5626), 1755–8. doi:10.1126/science.1082976

Schwartz, S., Maquet, P., & Frith, C. (2002). Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(26), 17137–42. doi:10.1073/pnas.242414599

Sergerie, K., Chochol, C., & Armony, J. L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *32*(4), 811–30. doi:10.1016/j.neubiorev.2007.12.002

Sevinc, G., & Spreng, R. N. (2014). Contextual and Perceptual Brain Processes Underlying Moral Cognition: A Quantitative Meta-Analysis of Moral Reasoning and Moral Emotions. *PLoS ONE*, *9*(2), e87427. doi:10.1371/journal.pone.0087427

Shin, L. M., Dougherty, D. D., Orr, S. P., Pitman, R. K., Lasko, M., Macklin, M. L., … Rauch, S. L. (2000). Activation of Anterior Paralimbic Structures during Guilt-Related Script-Driven Imagery.

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*(8), 334–40. doi:10.1016/j.tics.2009.05.001

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron*, *41*(4), 653–62.

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, *1156*, 81–96. doi:10.1111/j.1749-6632.2009.04418.x

Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., & Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *NeuroImage*, *23*(3), 967–74. doi:10.1016/j.neuroimage.2004.07.054

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *70*(6), 1256–69.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*, 345–72. doi:10.1146/annurev.psych.56.091103.070145

Taylor, J. L., & Novaco, R. W. (2013). A brief screening instrument for emotionally unstable and dissocial personality disorder in male offenders with intellectual disabilities. *Research in Developmental Disabilities*, *34*(1), 546–53. doi:10.1016/j.ridd.2012.09.018

Todd, R. M., & Anderson, A. K. (2009). Six degrees of separation: the amygdala regulates social behavior and perception. *Nature Neuroscience*, *12*(10), 1217–8. doi:10.1038/nn1009-1217

Ullsperger, M., & von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *The Journal of Neuroscience□: The Official Journal of the Society for Neuroscience*, *23*(10), 4308–14.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, *30*(3), 829–58. doi:10.1002/hbm.20547

Venkatraman, V., & Huettel, S. A. (2012). Strategic control in decision-making under uncertainty. *The European Journal of Neuroscience*, *35*(7), 1075–82. doi:10.1111/j.1460-9568.2012.08009.x

Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F., & Huettel, S. A. (2009). Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron*, *62*(4), 593–602. doi:10.1016/j.neuron.2009.04.007

Venkatraman, V., Rosati, A. G., Taren, A. a, & Huettel, S. a. (2009). Resolving response, decision, and strategic control: evidence for a functional topography in dorsomedial prefrontal cortex. *The Journal of Neuroscience□: The Official Journal of the Society for Neuroscience*, *29*(42), 13158–64. doi:10.1523/JNEUROSCI.2708-09.2009

Virani, K., Jesso, S., Kertesz, A., Mitchell, D., & Finger, E. (2013). Functional neural correlates of emotional expression processing deficits in behavioural variant frontotemporal dementia. *Journal of Psychiatry & Neuroscience□: JPN*, *38*(3), 174–82. doi:10.1503/jpn.120008

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*(3), 829–41.

Wagner, U., Handke, L., Dörfel, D., & Walter, H. (2012). An Experimental Decision-Making Paradigm to Distinguish Guilt and Regret and Their Self-Regulating Function via Loss Averse Choice Behavior. *Frontiers in Psychology*, *3*(October), 431. doi:10.3389/fpsyg.2012.00431

Wagner, U., N'Diaye, K., Ethofer, T., & Vuilleumier, P. (2011). Guilt-specific processing in the prefrontal cortex. *Cerebral Cortex (New York, N.Y.□: 1991)*, *21*(11), 2461–70. doi:10.1093/cercor/bhr016

Wahlund, K., Sorman, K., Gavazzeni, J., Fischer, H., & Kristiansson, M. (2010). Attenuated subjective ratings and skin conductance responses to neutral and negative pictures in nonpsychopathic mentally disordered offenders with various diagnoses. *Psychiatry Research*, *180*(1), 30–4. doi:10.1016/j.psychres.2009.09.009

Walton, M. E., Devlin, J. T., & Rushworth, M. F. S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, *7*(11), 1259–65. doi:10.1038/nn1339

Wang, X. (2011). The role of anticipated guilt in intentions to register as organ donors and to discuss organ donation with family. *Health Communication*, *26*(8), 683–90. doi:10.1080/10410236.2011.563350

Waytz, A., Zaki, J., & Mitchell, J. P. (2012). Response of dorsomedial prefrontal cortex predicts altruistic behavior. *The Journal of Neuroscience□: The Official Journal of the Society for Neuroscience*, *32*(22), 7646–50. doi:10.1523/JNEUROSCI.6193-11.2012

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*(3), 655–64.

Wiech, K., Kahane, G., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2013). Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. *Cognition*, *126*(3), 364–72. doi:10.1016/j.cognition.2012.11.002

Yang, T. T., Menon, V., Eliez, S., Blasey, C., White, C. D., Reid, A. J., … Reiss, A. L. (2002). Amygdalar activation associated with positive and negative facial expressions. *Neuroreport*, *13*(14), 1737–41.

Young, L., & Koenigs, M. (2007). Investigating emotion in moral cognition: a review of evidence from functional neuroimaging and neuropsychology. *British Medical Bulletin*, *84*, 69–79. doi:10.1093/bmb/ldm031

Appendices

Appendix A: Experiment Instructions

In this experiment, we are studying brain responses when humans make charitable donations. You will be shown a list of charities with a short description of each. While in the MRI scanner, you will be presented with individuals that are in need of help from these various charities and you will be given the opportunity to help to people like them by donating to these charities. You will have a limited amount of money to donate to the charities, thus please be sure to allocate the funds according to individuals you most want to support. For example, if you feel strongly towards one individual, you may choose to donate twice to the individual. In contrast, if you are sympathetic but are not particularly passionate about the individual's situation, you may choose not to donate. However, if you are sympathetic towards individuals but do not want to donate to them due to the high cost, you may be allowed to donate a lesser amount to them on your second choice. Remember that you will not be able to donate to all individuals that you will read about and will inevitability be unable to help some of them. Also, these individuals are in extreme need and therefore any amount of money donated will be highly useful to help them through their situations.

You will be presented with 26 scenarios of different individuals during each run of the experiment. You will complete 6 runs of the task. During the run, you will have two opportunities to donate to any given individual and you will be able donate to roughly 16 times in total (for approximately half of the trials). You will be paid a bonus for participating that will be a proportion of the money you choose not to donate.

There are two types of scenarios that you will encounter during the experiment. In the active scenario, after being presented with the situation of the individual, you will be prompted with the question of “Will you donate $10 for [name] today?”, where [name] will be the name of the individual that was introduced prior to the screen. You will then make a decision and push the button corresponding to that decision; 1 for donate and 2 for no donate. Also, every time you are asked to donate, you will have 3.5 seconds to make a decision. If you do not make a decision during this time, you will be prompted to respond faster.

In the passive scenario, you will be presented with a scenario again, but the prompt screen will this time say “Will a $10 donation be made for [Name] today?”. In these scenarios, an external source will decide whether a donation will or will not be made and you are only required to prompt the advancement of the screen by pressing either button 1 or 2. Keep in mind, however, that your button press will have no effect upon the subsequent decision that is made by the external source. This external source, however, will affect the amount of money you have available to donate and will thus affect your bonus.

Active scenario: “Will **you** donate $10 for [name] today?”

Passive scenario: Will **a** $10 donation be made for [Name] today?”

Once you have run out of money, you will be informed of this, and will be prompted to respond with the no donation button (button 2) for the remainder of the trials. While it is still possible press button 1, there will be no effect, and you will still be unable to donate any money. Finally, please remember that these individuals are in extreme need and therefore any amount of money donated will be a great help to them.

Appendix B: Subscale Means and Standard Deviations of Trait Scales for College Students

Guilt Inventory	Mean	SD	Low	High	Range
State Guilt	27.10	6.87	10	50	40
Moral Standards	47.21	7.93	23	71	48
Trait Guilt	54.96	12.33	23	97	74

TOSCA	Mean	SD	Low	High	Range
Shame	43.21	8.96	N/A	N/A	N/A
Guilt	61.83	6.84	N/A	N/A	N/A
Detached	31.84	5.49	N/A	N/A	N/A
Externalization	38.05	8.01	N/A	N/A	N/A
Alpha Pride	20.09	2.58	N/A	N/A	N/A
Beta Pride	20.8	2.72	N/A	N/A	N/A

PPI-R	Mean	SD	Low	High	Range
Machiavellian Egocentricity	44.98	8.99	N/A	N/A	N/A
Rebellious Nonconformity	33.59	7.83	N/A	N/A	N/A
Blame Externalization	30.62	7.45	N/A	N/A	N/A
Carefree Nonplanfulness	36.34	7.25	N/A	N/A	N/A
Social Influence	48.40	8.96	N/A	N/A	N/A
Fearlessness	38.24	8.73	N/A	N/A	N/A
Stress Immunity	35.80	7.22	N/A	N/A	N/A
Coldheartedness	33.10	6.79	N/A	N/A	N/A
Virtuous Responding	24.75	4.67	N/A	N/A	N/A
Deviant Responding	12.64	2.54	N/A	N/A	N/A
Total	301.06	31.26	N/A	N/A	N/A
Factor Self-Centred Impulsivity	N/A	N/A	N/A	N/A	N/A
Factor Fearless Dominance	N/A	N/A	N/A	N/A	N/A
Inconsistent Responding 15	8.49	3.12	N/A	N/A	N/A
SubIR40	N/A	N/A	N/A	N/A	N/A
Inconsistent Responding 40	2.29	6.50	N/A	N/A	N/A

Curriculum Vitae

AMBROSE TY

EDUCATION

2012 – Present Masters of Science Candidate
Neuroscience
Western University, Canada

2008 – 2012 Bachelor of Science
Neuroscience
University of Toronto, Canada

Research Grants

2012-Present Lawson Internal Research Fund Student Award

PUBLICATIONS

Greening, S., Norton, L., Virani, K., **Ty, A**., Mitchell, D., & Finger, E. (2013). Individual differences in the anterior insula are associated with the likelihood of financially helping versus harming others. *Cognitive, Affective & Behavioral Neuroscience*. doi:10.3758/s13415-013-0213-3

McCormick, C., St-Laurent, M., **Ty, A**., Valiante, T. a, & McAndrews, M. P. (2013). Functional and Effective Hippocampal-Neocortical Connectivity During Construction and Elaboration of Autobiographical Memory Retrieval. *Cerebral Cortex (New York, N.Y.: 1991)*. doi:10.1093/cercor/bht324

POSTER PRESENTATIONS

Differential Hippocampal Recruitment during Construction and Elaboration of Autobiographical Memories in patients with Temporal Lobe Epilepsy. Toronto Western Research Institute, Toronto, May 2012

An fMRI study on the neural correlates of guilt and restitution**.** London Health Sciences Centre, London, February 2013

Trait Guilt is associated with increased activity in ventrolateral and medial prefrontal cortex during acts of restitution, London Health Research Day, London, February 2014

Trait Guilt is associated with increased activity in ventrolateral and medial prefrontal cortex during acts of restitution, Cognitive Neuroscience Society, Boston, April 2014

Teaching Experience

2012-2013 Teaching Assistant, Introduction to Psychology (Psychology 1000), *Western University* Ontario, London, ON

Supervisors: Dr. Mike Atkinson

- Monitor online discussion boards, meet with students to help students with questions, host review sessions, mark assignments, and proctor exams