Behavioural and neurobiological foundations of vicarious processing

Patricia L. Lockwood

Division of Psychology and Language Sciences
University College London, UK

A thesis submitted for the degree of Doctor of Philosophy

Prepared under the supervision of:
Prof. Essi Viding & Prof. Jonathan P. Roiser

August 2015

Signed Declaration

I, Patricia L. Lockwood, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Empathy can be broadly defined as the ability to vicariously experience and to understand the affect of other people. This thesis will argue that such a capacity for vicarious processing is fundamental for successful social-cognitive ability and behaviour. To this end, four outstanding research questions regarding the behavioural and neural basis of empathy are addressed 1) can empathy be dissected into different components and do these components differentially explain individual differences in social functioning? [**Chapter 2**] 2) how does empathy relate to trait prosocial behaviour and do additional trait constructs moderate the association between empathy and prosocial behaviour? [**Chapter 3**] 3) how does the brain represent vicarious information, and does this vary dependent on individual differences in typical and atypical empathy? [**Chapters 4-5**] 4) what are the behavioural and neural mechanisms that link empathy to prosocial behaviour? [**Chapter 6**]

The findings of this thesis suggest that: 1) specific components of empathy have distinct associations with different kinds of disrupted trait social-cognitive ability 2) specific components of empathy are positively associated with trait prosocial behaviour and individual differences in the capacity to regulate one's own emotions moderates the strength with which empathy relates to trait prosocial behaviour 3) anterior cingulate cortex function may be critical in the perception of vicarious information, including pain and reward processing; and individual differences in anterior cingulate cortex function during pain and reward processing relates to individual differences in empathy and 4) empathy is important for prosocial decision making and underpins the neural computations that signal outcomes for others that are different from our expectations. Together, these findings contribute to a more complete and coherent understanding of the structure, function and neural basis of empathic/vicarious processing.

Acknowledgments

I would like to thank Essi for her wonderful supervision and support. I cannot thank her enough for her expertise, advice, encouragement and always finding time to meet with me. She has been the best mentor I could have hoped for and an inspirational role model. I offer my deepest gratitude to her and to Eamon for offering me an RA position that was the integral springboard to this whole process. I would also like to thank Jonathan Roiser for imparting his technical skills upon me and teaching me how to be a good researcher, Cat Sebastian for spending her invaluable time training me in the intricacies of fMRI and Geoff Bird for his expert advice and lively discussions.

I would like to say a special thank you to all of my colleagues and friends in Bedford way and the ICN–to my PhD buddies Phil, Lucy, Cha, Ana, Elif and Chloe who made my time in the PhD office fun, scholastic and immensely supportive. In particular to Lucy, who has been a special PhD buddy attending multiple conferences with me and always providing both emotional support and scholarly advice. Also to Steph, JB, Liz, Vanessa, Rachael and Ruth for their friendship and kindness.

Several people played a role in the collection and analysis of this data including my excellent confederates – Lewis Pollock and Jonathan Block. I am also very thankful to Vincent Valton who taught me computational modelling and Trevor, Marty, Joe and Prethra for helping me with setting up my fMRI experiments. Thank you also to Dave, Morgan and David for keeping my mac happy and always doing so with good humour and to Gill and John for keeping my finances in check.

I was fortunate enough during my PhD to meet someone who has been wonderfully supportive both as a partner and as a collaborator. To Matt, thank you for putting up with constant discussions about experimental design, analysis and results and for teaching me so much, I am extremely grateful. It is not an exaggeration to say the work in this thesis would not be the same without your input, guidance and excellent advice.

I am very grateful to all of my friends and family for their incredible support. To Harry for always being there for me and for being my psychology partner in crime, to Mills, Anna, Isobel and Megan for taking good care of me and keeping me sane, and to members of the Sunday Crèche and Norf london massive for their regular entertainment. I am most grateful to my mum for always believing in me and for making the whole process of this work so much easier through her care, support, good cooking and encouragement. Thank you to my Dad and Laurie for their support and for making me think beyond cognitive neuroscience. Finally, I would like to thank the MRC for funding this research and giving me the opportunity to pursue a career in academia.

Table of contents

List of figures

List of tables

List of manuscripts resulting from work in this thesis

Published/accepted

Lockwood, P. L. Apps, M. Roiser, J. Viding, E. (2015). Encoding of vicarious reward prediction in anterior cingulate cortex and relationship with trait empathy. *The Journal of Neuroscience*, 35(40):13720-13727.

Lockwood, P.L. (2015). Psychopathy: An Overview. *The Quarterly,* 94, 13-16.

Lockwood, P.L. Seara-Cardoso, A. & Viding, E. (2014). Emotion regulation moderates the association between empathy and prosocial behaviour. *PLoS One*. 9(5): e96555.

Lockwood, P.L. Sebastian, C. McCrory, E**.,** Hyde, Z. Gu, X. DeBrito, S. Viding, E. (2013). Association of callous traits with reduced neural responses to others' pain in children with conduct problems. *Current Biology,* 23 (10) 901-905.

Lockwood, P.L.. Bird, G. Bridge, M. & Viding E. (2013). Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Frontiers in Human Neuroscience*, 7:760.

In prep

Lockwood, P. L. Apps, M. Valton, V. Viding, E. Roiser, J. (in prep). Prosocial learning: vicarious reinforcement when choosing for others.

CHAPTER 1: General Introduction

1.1. Why should we care about social cognition?

Humans are highly social creatures, living in complex social environments and spending much of our lives interacting with, and thinking about, other people. In many species, aspects of social interaction are critical for individual fitness and well-being. For example, dominant male macaques have greater reproductive success than their submissive counterparts (Schülke, Bhagavatula, Vigilant, & Ostner, 2010). Humans and non-human primates with larger social networks have increased grey matter in brain regions involved in processing social information (reviewed in Rushworth, Mars, & Sallet, 2013). The complexity of social interactions individuals experience is a major determinant of variability in forebrain size, (Shultz & Dunbar, 2010). Taken together, these studies provide clear indication of the importance of social interaction for survival and evolutionary fitness (Silk, 2007).

Disturbances in social cognition and behaviour also characterise a number of psychiatric and neurological disorders, which have a large negative impact both on individuals and society as a whole (Kennedy & Adolphs, 2012). This thesis will argue that empathy, the ability to vicariously experience and to understand the affect of other people, is one of the key processes that can aid in successful social functioning and explains important variability in social cognition and behaviour.

1.2. What is empathy?

1.2.1. *Background, definitions and structure of empathy*

The psychologist Edward Titchener first introduced the word “empathy” into the English language over 100 years ago, as a translation of the German word *Einfühlung* ("feeling into"). Whilst there is no complete consensus as to the precise definition of empathy, most theorists agree that empathy is, broadly, the ability to vicariously experience and to understand the affect of other people (Bird & Viding, 2014; Decety & Jackson, 2004; Eisenberg, 2000; Hoffman, 2008; Singer & Lamm, 2009, but see Batson, 2009 for a different perspective). The ability to vicariously experience and to understand the affect of other people is the definition of empathy that will be adopted in the current thesis.

An important distinction within the structure of empathy is often made between emotional/affective [1] and cognitive aspects. Affective empathy is commonly understood as an affective state, caused by sharing the emotions of another person through observation or imagination of their experience (de Vignemont & Singer, 2006; Singer & Lamm, 2009). Although the observer’s emotional state is isomorphic with the other person’s state, the observer is aware that the other person is the source of their state (de Vignemont & Singer, 2006). Cognitive aspects of empathy are commonly referred to as perspective-taking, mentalising or theory of mind. These are those computations that enable the observer to understand another person’s beliefs and desires (Frith & Frith, 2006). It is important to note that some authors define empathy as comprised only of the “affective” components and label the “cognitive” components as a separate but related construct of “theory of mind” or “mentalising” on the basis that they rely on largely distinct neurocognitive circuits (e.g. Singer, 2006). In this thesis both components are seen as important contributors to the experience of empathy (in line with e.g. Bird & Viding, 2014). However, these

[1] These terms are used interchangeably in the current thesis.

different components, although positively associated, can be dissociated (see **Chapter 2**).

The origins of the distinction between affective and cognitive processes related to empathy can be traced back to historical accounts of social cognition regarding how we come to know the minds of others (Decety & Ickes, 2011). One account, known as theory-theory posited that we understand others' minds by forming a folk psychological theory, that is, a set of concepts about others (beliefs and desires) and governing principles as to how these concepts interact (e.g. people act to satisfy their desires according to their beliefs). A second account, known as simulation-theory was proposed as an alternative to theory-theory (Gordon, 1986; Heal, 1996). The core aspect of this alternative theory was that we understand the minds of others via a process simulation. The simulation-theory of social cognition has been hugely influential in the emerging field of social cognitive neuroscience. The discovery of 'mirror-neurons', neurons that increase their spike rate when a monkey makes a goal-directed action and also when the same actions are observed being performed by another, was invoked as evidence for simulation-theory (although see Hickok, (2014) and Cook, Bird, Catmur, Press, & Heyes, (2014) for critiques of the mirror-neuron theory). More recently, it has been generally accepted that both theory theory and simulation theory describe central aspects of social cognition that are needed to successfully understand and share another person's experience, and thus to empathise with them. However, it is important to keep in mind these two historical accounts, since they still influence what kind of profile a neural system should have for it to be involved in empathy (which will be discussed in further detail in the last section of this chapter, **1.5**).

It is also important to consider how cognitive and affective components of empathy relate to one another, as well as the sub-components that they comprise. It is generally agreed that affective empathy should be distinguished from emotion contagion, mimicry, empathic concern, compassion and sympathy (Bird & Viding, 2014; Singer & Lamm, 2009). Although these processes usually occur in similar contexts they are not the same as empathy. A recent model of empathy, entitled the self-to-other model of empathy (SOME; Bird & Viding, 2014) highlights that emotional contagion is a key precursor to empathy but

does not have to involve a distinction between self and other. Thus, although emotion contagion may be necessary for empathy, on its own it is not sufficient. Empathic concern, which is also called 'sympathy' or 'compassion,' involves 'feeling for' the other person (Singer & Lamm, 2009) and is associated with motivation to alleviate their suffering. This construct is thus often linked to prosocial motivation (Batson, 1998; Singer & Lamm, 2009). Although empathic concern is frequently equated with empathy, and empathy may lead to empathic concern, empathic concern does not necessarily involve any affect sharing and thus is not synonymous with affective empathy. Instead, empathic concern is an emotional response stemming from the apprehension or comprehension of another's emotional state or condition. This is not the same as what the other person is feeling (or is expected to feel) but consists of feelings of sorrow or concern for another person (Eisenberg, 2000).

1.3. Individual differences in empathy and their measurement

1.3.1. *Typical empathy*

It is well known that the capacity for empathy varies substantially between individuals (Bird & Viding, 2014; Blair, 2005) Various self-report measures have been developed to capture variability in empathic responding. One of the first of these measures, the Interpersonal Reactivity Index (Davis, 1983) has been hugely influential in the field of empathy research (Bernhardt & Singer, 2012). The IRI contains subscales measuring empathic concern, perspective-taking, personal distress and fantasy. The perspective taking and fantasy subscales are suggested to measure cognitive empathy, whereas the empathic concern and personal distress subscales are thought to assess affective empathy. However, difficulties in understanding the relationships among the scales have been demonstrated. Many studies have often omitted the fantasy scale in their analyses as the link with empathy is unclear. The personal distress subscale

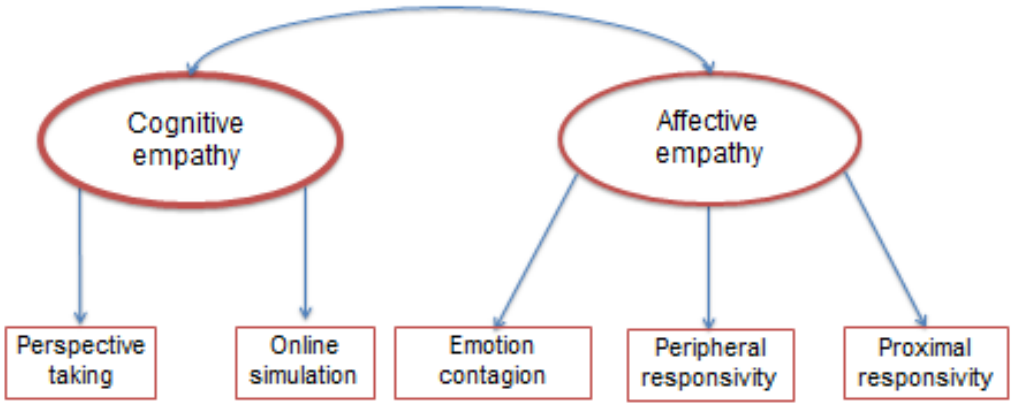
asks questions about personal responses to emergency situations, rather than empathy (Jolliffe & Farrington, 2006), and such responses may involve both empathising and sympathising (Jolliffe & Farrington, 2006). Moreover, the IRI possesses no specific measure of vicarious experience, only empathic concern, and thus does not measure the conceptualisation of empathy adopted in the current thesis and the field more generally.

To overcome these limitations and to create an instrument that assesses the multidimensional nature of empathy more closely, the Questionnaire of Cognitive and Affective Empathy (QCAE) was developed by Reniers and colleagues (Reniers, Corcoran, Drake, Shryane, & Völlm, 2011). The QCAE is an instrument devised to measure five key components of empathy. In the development of the QCAE, two raters selected items from other well-validated and commonly used empathy measures (e.g. Hogan Empathy Scale (HES; Hogan, 1969), Interpersonal Reactivity Index (IRI; Davis, 1983), Balanced Emotional Empathy Scale (BEES; Mehrabian & Epstein, 1972), and Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004) if they were deemed to measure empathy (see items below). Items deemed to measure other processes (e.g. sympathy) were not included. These items were then subjected to an exploratory factor analysis to identify the underlying structure of their associations and then to a confirmatory factor analysis in a separate sample to confirm the identified structure.

The five subscales that were identified by this procedure are: perspective-taking (e.g. ‘‘I can easily tell if someone else wants to enter a conversation.’’); online simulation (e.g. ‘‘Before criticizing somebody, I try to imagine how I would feel if I was in their place.’’); emotion contagion (e.g. ‘‘I am happy when I am with a cheerful group and sad when the others are glum.’’); peripheral responsivity (e.g. ‘‘I often get deeply involved with the feelings of a character in a film, play, or novel.’’); and proximal responsivity (e.g. ‘‘I often get emotionally involved with my friends’ problems’’). These subscales can be further grouped into two factors that the authors named cognitive and affective empathy. Cognitive empathy comprises the subscales of perspective-taking and online simulation, whereas affective empathy comprises the subscales of emotion contagion, peripheral responsivity and proximal responsivity (**see Figure 1.1**). The QCAE

has been shown to have well-validated psychometric properties (Reniers et al., 2011) and measures empathy as a multidimensional phenomenon comprised of related but separable constructs.

Figure 1.1. The two factor and five-factor solutions for the Questionnaire of Cognitive and Affective Empathy (QCAE, Reniers et al., 2011)

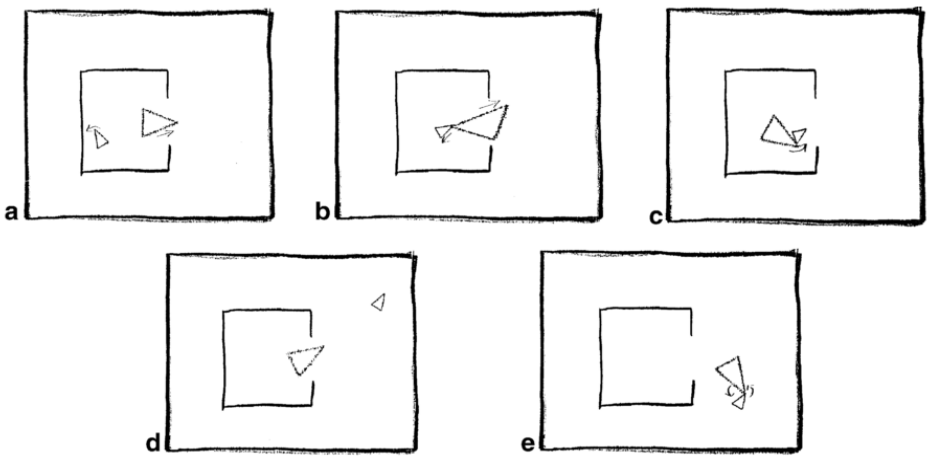


Given that the QCAE has been developed relatively recently there is a paucity of data regarding this measure and associations with external correlates, compared with what is available for the IRI. However, what has been done shows that the QCAE is likely to be capturing meaningful variance relating to different component processes of empathy and behavioural outcomes. For example, Seara-Cardoso, Dolberg, Neumann, Roiser, & Viding (2013) collected measures of psychopathic traits measured with the self-report psychopathy scale (SRP, to be discussed in the next section) and affective empathy, measured with the QCAE, in a large sample (n=100 women). They found that affective empathy was negatively associated with affective-interpersonal psychopathic traits (Seara-Cardoso et al., 2013). Yoder & Decety (2014) found that the cognitive empathy subscale of the QCAE was positively associated with ratings of blame in a moral judgment task. In neuroimaging studies, the perspective-taking subscale has been associated with neural response when punishing others in neural regions associated with processing mentalising

related computations (dorsomedial prefrontal cortex (dmPFC), posterior superior temporal sulcus (pSTS) (Molenberghs et al., 2014). Taken together these studies provide external validity for the QCAE. Where self-reported individual differences in empathy in adults are measured, in **Chapters 3** and **5** and **6**, the QCAE is adopted as it more precisely captures important facets of empathic processing than the IRI and more closely reflects the definition of empathy adopted in the current thesis.

A variety of behavioural tasks have also been developed that measure processes associated with empathy. To capture both typical and atypical variability in mentalising, a process that may contribute to cognitive empathy, one of the most common and well-validated tasks is the Theory of Mind animations task (Abell, Happe, & Frith, 2000).

Figure 1.2. Five stills taken from one of the animations scripted as Coaxing (mother and child).



Notes: (a) Mother tries to interest child in going outside. (b) Child is reluctant to go out. (c) Mother gently nudges child towards door. (d) Child explores outside. (e) Mother and child play happily together.

In this task participants' are required to understand others' complex mental states (e.g. tricking, coaxing). Each animation features two characters; a big red and small blue triangle either interacting with one another (ToM animations) or moving randomly (random animations). Participants are asked to watch each animation carefully and to describe what is happening whilst their verbal

responses are recorded. Participant responses are then rated by two people in terms of the use of language terms about intentionality and appropriateness. The intentionality scale ranges from 0 (no appreciation of another agent, nor actions or mental states) to 5 (the agent acts with the goal of affecting or manipulating the other agent's mental states). The appropriateness scale ranges from 0-3. This task has been previously used to examine ToM abilities in both typical and atypical populations (e.g. Abell et al., 2000; Castelli, Frith, Happé, & Frith, 2002; Jones, Happé, Gilbert, Burnett, & Viding, 2010) and it is sensitive enough to capture variability in understanding others intentions (See **Figure 1.2**).

In terms of the measurement of affective empathy there is a lack of empirical paradigms that directly measure vicarious experience. A recently developed behavioural measure of affective aspects of empathy (as well as cognitive aspects) is the Multifaceted Empathy Test (MET; Dziobek et al., 2008). In this task participants are presented with photo stimuli showing a naturalistic scene, such as a man standing in a kitchen looking sad. Participants make ratings in a number of separate stages. To measure "cognitive empathy" they are asked to label the emotion that the agent in the photo is feeling by selecting from 4 separate options. To measure "indirect" "affective empathy" (which the authors term emotional empathy) they are asked to rate how aroused they feel using the Self Assessment Manikin (SAM) scale, which goes from a very unaroused looking manikin to an aroused manikin. Finally, to assess "direct" "affective empathy" participants are asked to rate how concerned they feel for the person in the photo. Dziobek and colleagues found that adults with Asperger's syndrome were impaired in "cognitive empathy" but not "emotional empathy" as measured by this task. Whilst this task has some strengths, in that it can measure multiple components of empathy simultaneously, there are also limitations. The rating of calm/aroused for emotional empathy does not capture the condition of empathy that the emotional state is isomorphic with the agent being empathised with and measures emotional arousal rather than emotional valence. The explicit emotional empathy component measures empathic concern i.e. how concerned do you feel for the other person, rather than affect sharing, whilst the cognitive empathy component involves identification of

emotion rather than empathising. In order to develop a task that more closely probes vicarious experience, Seara-Cardoso et al., (2012) devised the SAM faces task (Seara-cardoso, Neumann, Roiser, Mccrory, & Viding, 2012) (see **Figure 1.3**).

Figure 1.3. Example trials from the Self-Assessment Manikin Faces task (SAM-Faces).

In this task participants are required to rate their *own* emotional response to the affective state of another on a 9-point manikin (ranging from smiling to a sad face with a neutral expression in the middle) whilst viewing images depicting a person showing either a sad, fearful, angry, happy or neutral expression. This task is thought to tap into the affective empathy construct as it not only estimates participants' vicarious response to emotional stimuli, but also comprises elements of self-awareness (participants have to evaluate their emotional response) and self/other distinction (participants are asked how the stimulus makes them feel, although note that like the MET this task does not measure an emotional response as isomorphic). In **Chapter 2** the ToM animations and SAM task are used to explore the structure of empathy.

1.3.2. *Atypical empathy*

Disturbances in empathy exist in a number of psychiatric and neurological disorders. Studying individuals with disorders of empathy can allow us to explore the behavioural and neural basis of empathy, as well as, hopefully, to help individuals with these disorders. In the next section I focus on two prominent disorders of social abilities, psychopathy and autism. Those who suffer from these disorders behave in ways that suggest that they may lack empathy. However, their seeming lack of empathy is likely to occur for different reasons.

1.3.2.1. Psychopathy

Psychopathy is a disorder characterised by a constellation of cognitive and behavioural atypicalities including callousness, shallow affect, lack of guilt, antisocial behaviour and impulsivity (e.g. Blair, Mitchell, & Blair, 2005; Blair, 2013; Cleckley, 1941; Hare, 1999; Hare & Neumann, 2006). These individuals commit a disproportionate amount of violent crime, and place a substantial economic and emotional burden on society (Anderson & Kiehl, 2012). The ability of individuals with psychopathy to seriously violate the rights of others' is thought to highlight a disturbance in an appropriate empathic response to other people (Blair, 2005), thus psychopathy is perhaps the archetypal empathy disorder. Adults with psychopathy show reduced affective response to the distress of others (Blair, Jones, Clark, & Smith, 1997), blunted emotional reactivity to aversive stimuli (Levenston, Patrick, Bradley, & Lang, 2000), impaired recognition of distress cues in others (Blair, Colledge, Murray, & Mitchell, 2001) and perhaps also positive facial expressions, (e.g. Brook, Brieman, & Kosson, 2013), and atypical neural responses to stimuli depicting others experiencing pain (Decety, Skelly, & Kiehl, 2013; Meffert, Gazzola, Boer, Bartels, & Keysers, 2013). Similarly, adults with high levels of psychopathic

traits seem to show reduced affective resonance with processing of both positive and negative emotions (Seara-Cardoso et al., 2013; Seara-cardoso et al., 2012). They also report less enjoyment of interacting prosocially with others, suggesting reduced vicarious experience more broadly (Foulkes, McCrory, Neumann, & Viding, 2014).

By contrast, one of the defining features of psychopathy is the ability to successfully manipulate others (Hare, 1999), a process that requires the ability to understand another person's thoughts. Thus it might be expected that individuals with psychopathy would have intact theory of mind and may even be able to compute other's likely emotional states, even if they do not share such states themselves (i.e. they may have intact "cognitive empathy"). Several studies report no "cognitive empathy" impairments (Blair et al., 1996; Dolan & Fullam, 2004; Richell et al., 2003) and even superior ability (Hansen, Johnsen, Hart, Waage, & Thayer, 2008) in individuals with psychopathy or high psychopathic traits. Some studies have reported problems in tasks that the authors have claimed to asses "cognitive empathy" in both incarcerated psychopaths (Brook & Kosson, 2013) and healthy samples with high psychopathic traits (Ali & Chamorro-Premuzic, 2010). One possibility for these mixed findings is that different 'cognitive empathy' paradigms vary in their level of affective content, with some measures requiring identification of other people's feelings, rather than simply processing others' intentions. Those studies that have reported a negative association between "cognitive empathy" and psychopathy/high psychopathic traits have all involved emotional content, and it is currently not clear whether the findings in these studies are driven by problems related to basic affective processing, rather than difficulties in "cognitive empathy" per se.

In children, there is abundant evidence that psychopathic traits and behaviours can be observed and that the behavioural and affective disturbances that are seen generally mirror those observed in adults with high levels of psychopathic traits. In childhood, high levels of antisocial behaviour can be diagnosed as conduct disorder (DSM-5). Particular subsets of children with conduct disorder can also have elevated levels of psychopathic traits, which are termed callous-unemotional traits in research studies and "limited prosocial emotions" in the

new DSM-5 guidelines. Evidence of a similar profile of empathic impairments in children with CU traits as compared to adults with psychopathy or high psychopathic traits was reported by Jones et al. and Schwenck et al. (Jones et al., 2010; Schwenck et al., 2012). These authors found that children with psychopathic traits showed less affective resonance (affective empathy) with others' emotions but did not have problems with cognitive perspective-taking (cognitive empathy). Callous-unemotional traits in children can persist into adulthood (Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007) and are highly heritable (Viding, Blair, Moffitt, & Plomin, 2005). In contrast, antisocial behaviour in children without callous-unemotional traits appears to be primarily driven by environmental influences and is typically less persistent (Frick, Ray, Thornton, & Kahn, 2014; Viding et al., 2005).

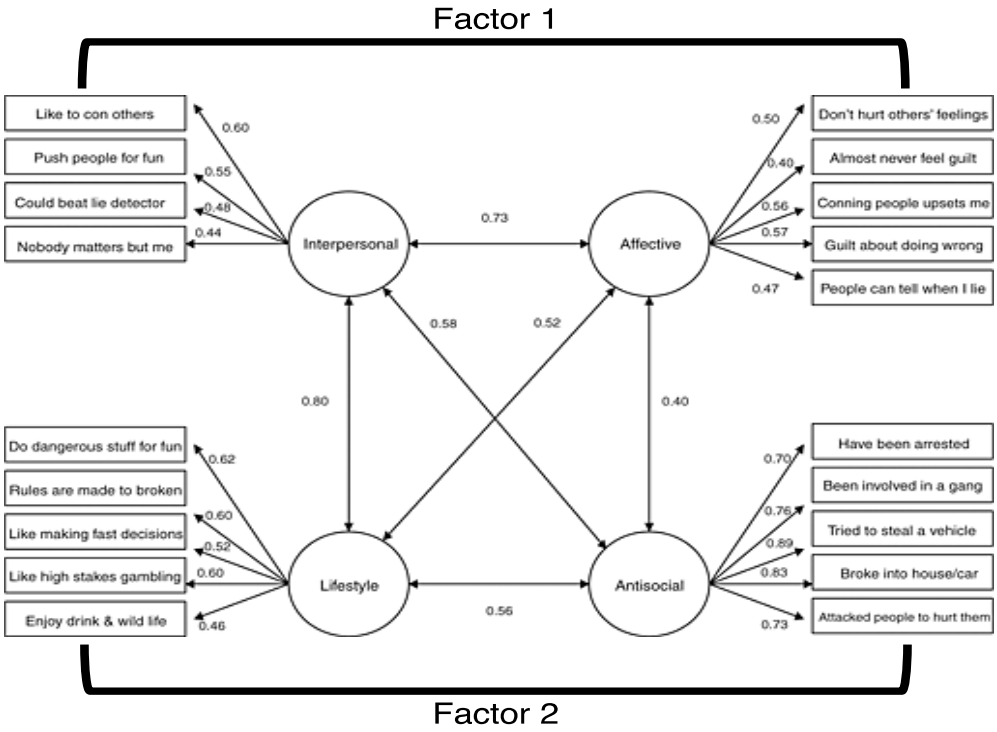
In terms of the pathways through which psychopathy develops, researchers have proposed that individuals with psychopathy have an atypical experience of distress, such as fear or sadness (Blair, 2013) underpinned by dysfunction in specific neural systems. Genetic and environmental factors influence the development of these neural systems. Throughout development, the reduced ability to experience emotions results in impaired associations between antisocial actions and outcomes of causing distress in other people (Bird & Viding, 2014; Blair, 2013), and also perhaps outcomes of causing positive experiences in others (e.g. Brook et al., 2013). Reduced distress in an infant also results in fewer opportunities in the environment for learning which cues reliably signal distress in other people (Bird & Viding, 2014; Blair, 2013). Researchers have argued that it is the reciprocal interaction between atypical emotional reactivity and the resulting interactions with the environment that can lead to the development of psychopathy (Bird & Viding, 2014).

Measuring psychopathic traits in adults and children

In forensic settings, the most widely used and validated instrument for assessing psychopathy is the Hare Psychopathy Checklist Revised (PCL-R;

(Hare, 1999). The PCL-R conceptualises psychopathy as consisting of two broad dimensions, termed Factor 1 and Factor 2 in the literature. Factor 1 comprises affective and interpersonal features including reduced empathy and guilt, as well as manipulation of others. Factor 2 comprises antisocial behaviour and impulsive lifestyle choices (Hare, 1999).

Figure 1.4. Four-factor model of psychopathy. Labels of Factor 1 and Factor 2 show conceptual and empirical overlap with the PCL-R.



Notes: Modified from Neumann et al. (2012) with permission of the copyright owner.

Psychopathic traits can be reliably measured in a typical adult population, with increasing evidence that these traits existing on a continuum (See Hare & Neumann (2008) for a review). Self-report measures suitable for non-forensic samples include the Self-Report Psychopathy Scale (Paulhus, Neumann, & Hare, 2015), which is perhaps the most well validated measure of psychopathic traits. The SRP has been shown to have a clear latent structure, which mirrors the factor structure of the PCL-R (Carré et al., 2013). The SRP is strongly

positively correlated with the PCL-R (Neumann & Pardini, 2014; Neumann, Schmitt, Carter, Embley, & Hare, 2012; Paulhus et al., 2015) and is associated with related external correlated including criminal offenses and externalizing psychopathology (Neumann & Pardini, 2014). There is evidence that external correlates of psychopathic traits in community samples mirror those observed in forensic samples in both behavioural and neural profiles (e.g. Seara-Cardoso & Viding., 2014), Where psychopathic traits in adults are measured in **Chapter 2** the SRP measure is adopted.

In children, conduct problems are generally assessed using the DSM criteria, as specified for example in the Child and Adolescent Symptom Inventory (CASI-4R; Gadow & Sprafkin, 2009) Conduct Disorder scale (CASI-CD). A research diagnosis of conduct disorder can be used where a child meets CASI-CD criteria for scores associated with a clinical diagnosis of CD (Gadow & Sprafkin, 2009). Callous-unemotional traits can be assessed using the Inventory of Callous-Unemotional Traits (ICU) (Essau, Sasagawa, & Frick, 2006). The ICU has been validated in a large sample (n=1443) of adolescents and contains subscales of callous, uncaring and unemotional traits. Fit indices suggest that callous-unemotional traits can be represented as this three-factor structure (callous, unemotional, uncaring) as well as a single higher order factor (callous-unemotional traits). In further studies, this three-factor structure was confirmed in a sample of juvenile offenders, and the ICU was documented to related to key external correlates including increased aggression, delinquency and psychophysiological and self-report indices of emotional reactivity (Kimonis et al., 2008). Thus, where CU traits are measures in children with conduct problems (**Chapter 4**) the ICU is used.

1.3.2.2. Autism

Autism spectrum disorders (ASD) refer to a class of developmental disorders characterised by impaired social and communication skills and a restricted repertoire of interests and activities (American Psychiatric Association, 2013).

Several decades of research indicate that individuals with ASD have difficulties with “cognitive” aspects of empathy (see Hill & Frith, 2003). ASD has often been described as a disorder associated with “poor empathy” (Baron-Cohen, 2011). However, it is important to note that the nature of their social information processing deficits and behaviours seem very different from those seen in individuals with psychopathy/psychopathic traits (Viding, McCrory, & Seara-Cardoso, 2014).

A number of studies measuring cognitive and affective processes related to empathy have found impairments in cognitive perspective-taking but not empathic concern in adults (Dziobek et al., 2008) and reduced cognitive perspective-taking (cognitive empathy) but not affective resonance/affective empathy in children, with ASD (Jones et al., 2010; Schwenck et al., 2012). Studies focusing solely on affective processing have found evidence of preserved affective processing, including normal skin conductance response to others’ negative emotions when emotions are unambiguous and presented under conditions of low distraction (Blair, 1999). Some theorists have argued that affective empathy is actually heightened in individuals with ASD (Smith, 2009) and reports of greater empathic facial affect in children with ASD compared to controls supports this (Capps, Kasari, Yirmiya, & Sigman, 1993).

However, individuals with ASD have also been found to have lower scores on the empathy quotient, a self-report questionnaire of empathy, compared to typically developing individuals (Baron-Cohen & Wheelwright, 2004). Another study found that parents of children with ASD reported their children to be less concerned about emotional situations and less responsive to distress cues than control children (Hudry & Slaughter, 2009). Nevertheless, it is unclear in studies that do find affective empathy impairments whether these relate to problems in social responsivity rather than affective empathy per se.

A further consideration for the profile of empathy in ASD is the high comorbidity of the disorder with alexithymia. Alexithymia is a sub-clinical condition defined by an inability to identify and describe one’s own feelings. Preliminary behavioural and neuroimaging research suggests that when genuine affective and empathy impairments are seen in individuals with ASD, these may be a

function of interoceptive difficulties related to alexithymia rather than ASD per se (Bird et al., 2010; Silani et al., 2008). After accounting for the variance explained by alexithymia in the studies assessing empathy in individuals with ASD, there is typically no difference in empathy between individuals with ASD and controls (Bird & Cook, 2013). However, one recent fMRI study found no significant moderating effects of alexithymia in a task where participants viewed pictures of other people in pain (termed an 'empathy for pain' task by the authors) in individuals with ASD (Fan, Chen, Chen, Decety, & Cheng, 2014). Nevertheless, the variance in alexithymia scores was very limited (SD 3.8 in (Fan et al., 2014) vs. 11.8 in (Bird & Cook, 2013), which may explain why no effect of alexithymia was observed.

To capture individual differences in autistic traits in community samples the Autism Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) is one of the most commonly used and best validated measures (e.g. (Chakraborty & Chakrabarti, 2015; Haffey, Press, O'Connell, & Chakrabarti, 2013). The AQ correlates highly with other measures of ASD severity (e.g. ADOS; Brugha et al., 2012). Consequently, where autistic traits are measured in **Chapter 2** the AQ is used, although it is acknowledged that high levels of autistic traits as measured by AQ do not equate to an ASD diagnosis.

1.4. Empathy and relationship with social behavior

1.4.1. *Antisocial and Prosocial behavior*

Associations between empathy and antisocial behaviour have mostly been investigated in children with conduct problems and in individuals with high levels of psychopathic traits. As highlighted in the previous section, theorists have suggested that antisocial behaviour in part occurs due to an atypical response to the suffering of others (Blair, 2005) and it is clear that adults with

psychopathy and children with psychopathic traits have high levels of antisocial behaviour as well as empathy impairments (Anderson & Kiehl, 2012). On the flip side, empathy is thought to be a crucial motivation factor for prosocial behaviour (Eisenberg, Eggum, & Di Giunta, 2010; Hoffman, 2008; Penner, Dovidio, Piliavin, & Schroeder, 2005). Prosocial behaviours can be broadly defined as social behaviours intended to benefit other people. Humans are thought to have a remarkable capacity to engage in prosocial behaviours with genetically unrelated individuals (Fehr & Fischbacher, 2003). People routinely engage in charitable donation behaviour and exhibit social preferences, which are influenced by positive or negative concern for the welfare of others (Fehr & Camerer, 2007).

Processes related to both affective and cognitive empathy are thought to be positively associated with prosocial behaviour (for a review see Eisenberg et al., 2010). In the majority of these studies the IRI questionnaire (Davis, 1983) and cardiovascular and electrodermal indices, such as heart rate deceleration and facial electromyography (EMG), have been used as proxy measures of affective empathy. For example, heart rate deceleration (which is thought to index vicariously induced sadness or sympathy (e.g. Eisenberg, McCreath, & Ahn, 1988) and increased indicators of facial sadness when watching needy others, are associated with increased willingness to help (Eisenberg et al., 1989). Dispositional empathic concern, as measured by the IRI, has also been linked to higher levels of self-reported charitable giving (Davis, 1983) and greater self-reported concern for the welfare of others (Batson, 1998).

In terms of associations between cognitive components of empathy and prosocial behaviour, the majority of studies have focused on correlating the perspective-taking subscale of the IRI to self-reported prosocial behaviour and have found that trait perspective taking is positively associated with frequency of volunteering (Carlo, Allen, & Buhman, 1999) and self-reported prosocial tendencies (Carlo, Hausmann, Christiansen, & Randall, 2003). As mentioned in section (**1.3.1**), however, the empathic concern and perspective taking scales of the IRI tap constructs that, although related, are different from the current conceptualisation of “affective empathy” and “cognitive empathy”. Consequently, there is a lack of empirical evidence that empathy is indeed a

motivating factor for prosocial behaviour. Importantly, the majority of studies suggesting empathy as a motivating factor for prosocial behaviour have investigated self-reported empathic concern (feeling 'for' another person, including compassion and sympathy, e.g.(Batson, 1998; Davis, 1983), rather than self-reported affective empathic responses (the ability to vicariously experience the emotional experience of others; or feeling 'as' another individual). While these two processes are no doubt closely related, there is a lack of empirical data regarding how feeling in a similar emotional state to another may motivate prosocial behaviour. In addition, self-reported cognitive empathic ability (i.e. the ability to position oneself 'in another person's shoes') might also relate to prosocial behaviour, but compared to the role of affective empathic processes motivating empathy this has received relatively little attention to date (except for work by(Carlo et al., 1999, 2003) that has associated perspective-taking as measured by the IRI with prosocial tendencies).

1.5. Neural underpinnings of empathy and social decision-making behaviour

In order to vicariously process the experiences of another person it is first necessary to perceive or comprehend their experience. This perception can then guide subsequent decision-making behaviour, linking empathy to social behaviour. In the next section I review studies that have attempted to uncover the neural basis of vicarious perception and vicarious decision-making and identify research questions that are still outstanding.

1.5.1. *Vicarious perception*

Studies in the field of social neuroscience have attempted to identify the neural substrate of vicarious perception, namely the perception of the experience of another person, which can perhaps be seen as a proxy measure of empathy. In general, this research has followed two somewhat separate strands. One strand has emphasised that social information is processed by a specific set of neural circuits referred to as the "social brain" (Behrens, Hunt, & Rushworth, 2009) that do not overlap with the processing of self-relevant information. This strand can be seen as historically influenced by theory-theory accounts of social cognition, and the brain areas implicated are in some cases is referred to as the 'mentalising' system. The second strand has instead focused on regions that show overlap between first-person and third-person processing of information, a natural derivation from simulation-theory accounts of social cognition. In general, some researchers refer to this as the 'mirroring system'.

Many studies investigating the neural basis of empathy have focused on the observation of other people's negative experiences. A seminal study by Tania Singer and colleagues, now with over 2500 citations, was one of the first to investigate the neural underpinnings of the observation of other people's negative experiences, as a proxy measure of empathy (Singer et al., 2004). In this study, participants experienced a painful stimulus whilst undergoing functional magnetic resonance imaging (fMRI). On "other" trials participants observed cues that signalled that their partner, who was present in the same room, was receiving a painful stimulus. Singer and colleagues found that the anterior insula (AI) and anterior cingulate cortex (ACC) responded both when the subject themselves received the painful stimulus and when they viewed a cue that indicated that their partner received a painful stimulus (though it is important to note the authors used a block design so they were not able to separate out cue and outcome related neural responses). In contrast, response in the secondary somatosensory cortex and primary somatosensory cortex was associated with greater response to the pain participants received themselves, compared to not receiving pain themselves. Trait empathic concern, as measured by the IRI, was positively associated with blood oxygen level (BOLD) response for the pain-no pain for other contrast. Taken together, these findings were the first to suggest that the observation of others experiences activates

similar neural regions to one's own experiences, which was interpreted as a neural marker of empathy. Since this seminal study, many studies have replicated the finding of AI and ACC response to the observation of others' negative experiences, particularly pain (reviewed in Lamm, Decety, & Singer, 2011), but also disgust (Jabbi, Bastiaansen, & Keysers, 2008; Wicker et al., 2003).

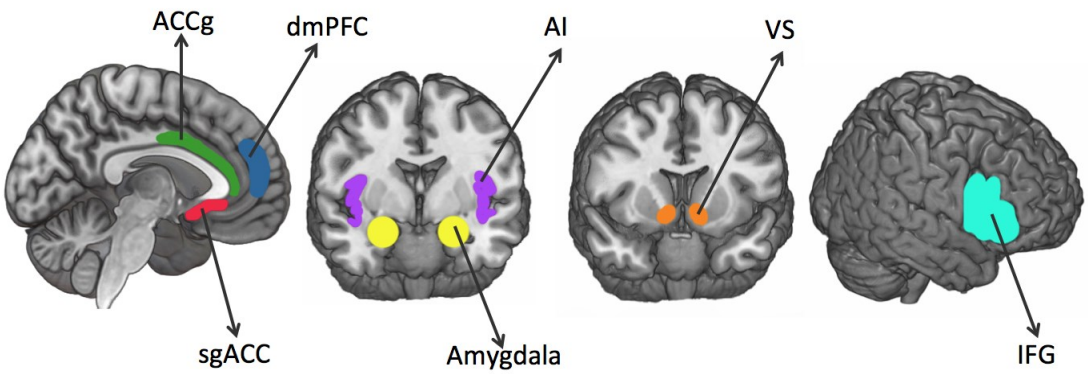
More recently paradigms have been used to examine response to other people's positive experiences, such as rewards (Apps & Ramnani, 2014; Braams et al., 2014; Mobbs et al., 2009) and pleasant touch (Lamm, Silani, & Singer, 2015). One of the first paradigms to examine neural responses to the observation of others' rewards was conducted by Mobbs and colleagues (Mobbs et al., 2009). In their study, participants watched videos of two game show contestants answering questions about their political and social views in a way that was socially desirable (SD) or socially undesirable (SU). Participants were then scanned whilst they watched these two players (SD and SU) play a card-guessing game where they could win or lose money. After, the participants played the game themselves. Mobbs et al. found that activity in the ventral striatum, subgenual cingulate cortex (sgACC) and ventromedial prefrontal cortex (vmPFC) correlated with the difference between watching the socially desirable and socially undesirable contestant win. Only the ventral striatum response was also observed when the participants played themselves. However, Mobbs and colleagues also used a block design so were unable to separate neural responses to reward prediction and reward consummation.

More recently, studies have focused on separating neural responses to cues and outcomes. This is of particular importance as there is evidence that different neural regions may process cues signalling reward, and the actual receipt of reward, for oneself (Rademacher et al., 2010). Apps et al., (2014) examined activity at the time of cues that signalled the net-value (benefit-cost) of anticipated reward magnitude (benefit) to be gained and the level of effort (cost) to be incurred either by a participant themselves or by a social confederate (Apps & Ramnani, 2014). They observed that the gyral portion of the ACC (ACCg) signalled the net-value of rewards gained by others. In contrast, activity in the ventral striatum signalled the net-value only for the

participants themselves. This suggests that whilst the ACCg may play a specific role in vicarious reward.

Meta-analysis of the observation of others pain (Lamm et al., 2011) and the observation of others positive outcomes (Morelli, Sacchet, & Zaki, 2015) suggest that the neural processing of other peoples experiences, in general, recruits the AI, the inferior frontal gyrus (IFG), ACC, subgenual ACC, medial prefrontal cortex, amygdala and ventral striatum (see **Figure 1.5**). The meta-analysis by Lamm et al., identified these regions as the same regions that also respond to first-person information in the same task. This fits with the shared representation hypothesis of empathy, namely that empathy relies on shared neural systems between first-person and third-person experience (which can be seen as influenced by simulation theory accounts of social cognition) (Engen & Singer, 2013). However, the meta-analysis on neural responses to others rewarding outcomes (including food and money) identified regions involved in processing both personal and vicarious reward but also regions that show distinct contributions to these processes (Morelli et al., 2015).

Figure 1.5. Brain regions that respond in paradigms of vicarious perception of other people's experience.

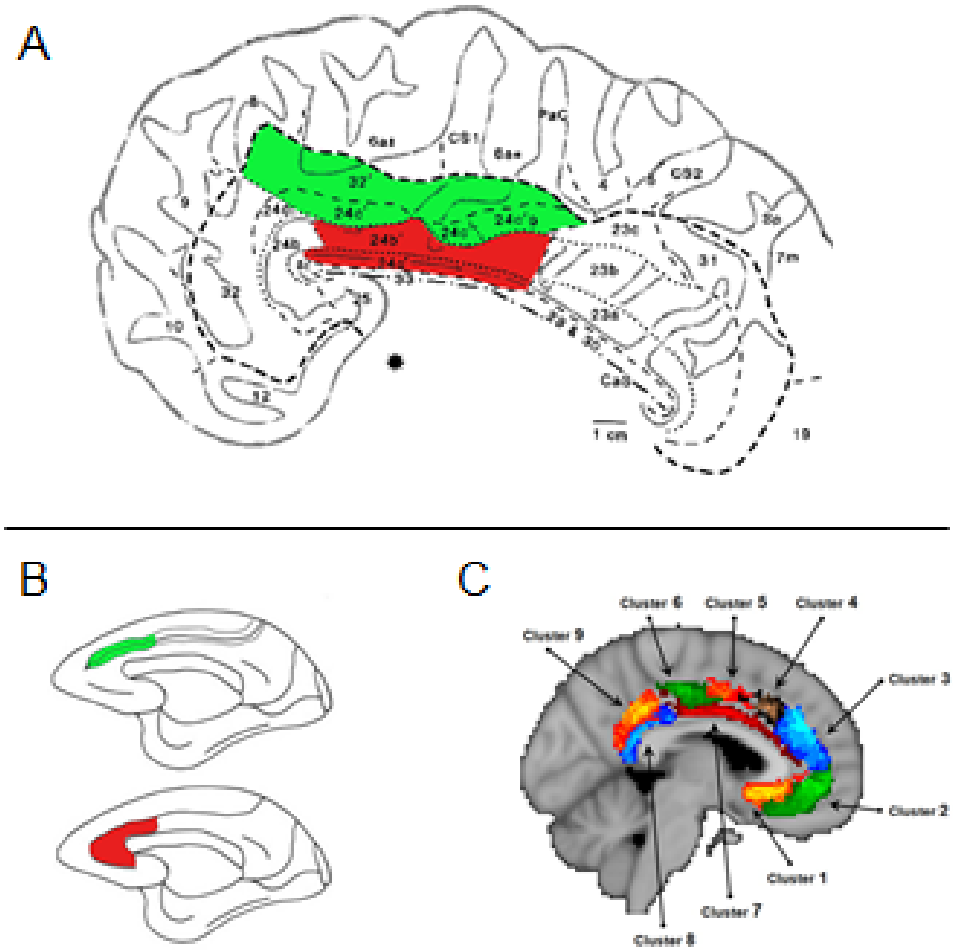


Notes: ACCg: gyral portion of the anterior cingulate cortex. dmPFC: dorsomedial prefrontal cortex. sgACC: subgenual portion of the anterior cingulate cortex. AI: anterior insula. VS: ventral striatum. IFG: inferior frontal gyrus.

Research in non-human animals can help us to identify neural regions that may be involved in empathising, as it allows us to directly record the activity of neurons during vicarious perception, and to cause focal lesions in an ethical manner (although of course we cannot infer from these studies that animals are experiencing the phenomenon of empathy). Of particular relevance, this research has identified important divisions within the ACC that appear to be central for understanding social behaviour. The cingulate cortex consists of five zones: retrosplenial, posterior (PCC), mid (MCC), anterior (ACC) and perigenual (including the sgACC) (Palomero-Gallagher, Mohlberg, Zilles, & Vogt, 2008; Vogt, Nimchinsky, Vogt, & Hof, 1995). In both the MCC and ACC there are differences in cytoarchitecture and connectivity between the sulcus (ACCs) and the gyrus (ACCg) that are indicative of distinct functional properties. There is converging evidence that the ACCg, as compared to the ACCs, plays a key role in social cognition and behaviour in both human and non-human primates (Apps, Green, & Ramnani, 2013; Apps & Ramnani, 2014; Behrens, Hunt, Woolrich, & Rushworth, 2008; Boorman, O'Doherty, Adolphs, & Rangel, 2013; Chang, Gariépy, & Platt, 2013; Jones et al., 2011; Rudebeck, Buckley, Walton, & Rushworth, 2006, and see **Figure 1.6**). It has been argued that the ACCg and ACCs both processes information that conforms to the principles of reinforcement learning theory (to be discussed in further detail in section **1.7**) but the ACCg does this in social contexts whilst the ACCs does so in 'non-social' contexts (Apps, Lockwood, & Balsters, 2013).

The claim of a key role for the ACCg in social cognition and behaviour is built upon several lines of evidence. Lesions to this region have been shown to impair the processing of social stimuli and cause a reduction in the execution of social behaviours (Rudebeck et al., 2006). Rudebeck and colleagues examined the latency that occurs before reaching for a food item when the food was presented simultaneously with different types of social and emotional stimuli. Monkeys without lesions, with lesions to the lateral OFC or with lesions to the ACC sulcus, showed a similar latency before reaching for a food item when it was presented with a social stimulus. In contrast, monkeys with a lesion specifically to the ACCg showed significantly decreased latencies in the presence of multiple different types of social stimuli.

Figure 1.6. The anterior midcingulate cortex (ACC/MCC).



Notes: (**A**) Cytoarchitecture of the MCC taken from Vogt et al., (1995). The areas shaded in green lie in the MCCs. The areas shaded in red lie on the MCCg. We argue that this area is engaged when processing information about others' decisions. Specifically we argue that areas 24a' and 24b', which lie on gyral surface of the cingulate cortex, extending on average 22mm posterior and 30mm anterior the anterior commisure. Lesion site of the MCCg and ACCg (**B**) and the MCCs and the ACCs from Rudebeck et al., (2006). The lesions that affected the gyrus caused disruptions to social behaviour and disrupted the processing of social stimuli. (**C**) Subdivisions of the MCC and ACC according resting-state connectivity (Beckmann et al., 2009). The cluster shown in dark red corresponds, broadly, to the ACCg/MCCg. Reproduced from Apps, Lockwood and Balsters (2013).

The latencies of the ACCg lesion monkeys were the same as the responses of control monkeys when the food was presented at the same time as a neutral stimulus, suggesting that the stimuli no longer held the same social significance.

ACCg lesions also resulted in a reduction in communicative behaviours such as lip smacking and communicative vocalisations. Importantly, in the monkeys with lesions of the OFC and ACCs no impairments in communicative behaviours or differences in response latencies to social stimuli were observed. This study therefore supports the notion that the ACCg processes social information and guides behaviours during social interactions.

A seminal study by Chang and colleagues was the first to record from neurons in the ACCs, ACCg and OFC during a social decision-making task (Chang et al., 2013). They observed that a greater proportion of neurons in the ACCg, compared to the ACCs and OFC, responded to cues that predicted rewards for other monkeys and also to decisions to allocate rewards to other monkeys. The ACCg is densely connected to regions that process social information, but also, to regions that process reward-related information (Haber, Kunishio, Mizobuchi, & Lynd-Balta, 1995; Lynd-Balta & Haber, 1994; Williams & Goldman-Rakic, 1998; Yeterian & Pandya, 1991). Finally, it has been found that in non-human primates those with larger social networks have increased grey matter volume in the ACCg compared to those with smaller social networks (Sallet et al., 2011). Taken together these studies provide support for the claim that the ACCg is important for processing others' rewards and also more widely in social behaviour.

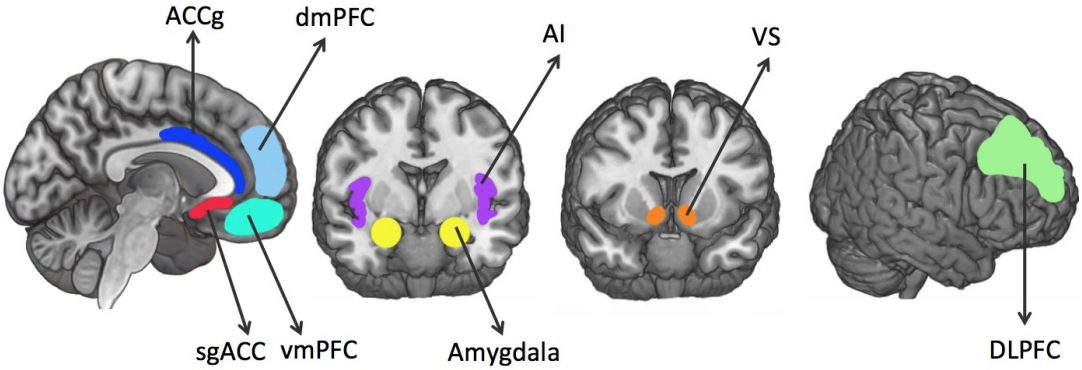
Overall, studies on vicarious perception have identified a network of regions that could be involved in encoding the experiences of other people. Whether these are the same regions involved in encoding first-person-experience and whether this is a necessary condition for empathy remains to be resolved. Meta-analyses of the perception of other peoples negative and positive experiences have, in general, identified the ACC/MCC, AI, sgACC, IFG and ventral striatum as potential neural regions involved in vicarious perception (Fan, Duncan, de Greck, & Northoff, 2011; Lamm et al., 2011; Morelli et al., 2015). Studies in non-human primates have suggested that the ACCg in particular may be key for processing others experiences and more widely in social behaviour. In the next section the role of these regions and others will be reviewed in relation to how the perception of others experiences may guide our social behaviour.

1.5.2. *Vicarious decision-making behavior*

Over more than a decade, research in the field of behavioural neuroscience has provided a rich theoretical characterisation of the behavioural and neural mechanisms that underpin value-based decision-making (Behrens et al., 2009). Recently, this framework has begun to be applied to decision-making in social environments (see Rilling & Sanfey, 2011; Ruff & Fehr, 2014 for reviews). Relevant contexts in which the ability to vicariously process another's experiences could be important for guiding our behaviour might include reward learning (Behrens et al., 2009), prosocial learning, observational learning (Burke, Tobler, Baddeley, & Schultz, 2010), and teaching (Apps, Lesage, & Ramnani, 2015). These same processes have been argued to be crucial for understanding social cognition and behaviour across species (Frith & Frith, 2012). As in research on vicarious perception, research on the neurobiological basis of social decision-making has also followed two somewhat separate approaches. One focusing on neural regions of 'socially-specific' cognition and one on the 'common-currency' of overlapping regions involved in social and non-social decision-making processes (Ruff & Fehr, 2014).

A number of neural regions have been linked to processing decisions in social environments. These regions include the ventral striatum, ACCg, sgACC, amygdala, insula, OFC, VMPFC, DMPFC and DLPFC (see Moll & Schulkin, 2009; Rilling & Sanfey, 2011; Ruff & Fehr, 2014 for reviews and **Figure 1.7**).

Figure 1.7. Brain regions that respond in vicarious decision-making paradigms.



Notes: Brain regions that are response in paradigms of vicarious decision-making paradigms. ACCg: gyral portion of the anterior cingulate cortex. dmPFC: dorsomedial prefrontal cortex. sgACC: subgenual portion of the anterior cingulate cortex. AI: anterior insula. VS: ventral striatum. DLPFC: dorsolateral prefrontal cortex.

As mentioned previously, in non-human primates the ACCg was identified as containing neurons which respond to the decision of a monkey to allocate a reward to another monkey (Chang et al., 2013). A recent study also identified neurons in the dorsal ACC (unclear whether this area was ACCg or ACCs) during a prisoners dilemma game (Haroush & Williams, 2015). In this latter study, the authors observed that stimulation of neurons in the dorsal ACC made monkeys more competitive.

In humans, one of the first studies to examine the neural correlates of rewards allocated to others observed that whilst the ventral striatum signalled both monetary reward and charitable donation, the sgACC showed a greater response for decisions to donate money to charity vs. monetary reward (Moll et al., 2006). This suggests support for both the common currency and socially-specific hypotheses of social decision-making. Whilst the ventral striatum signals both monetary reward and charitable donation, the sgACC region could be specifically involved in processing information about delivering rewards to others. Another neural region consistently identified in social decision-making

studies is the dorsolateral prefrontal cortex (DLPFC) (see Fumagalli & Priori, 2012; Young & Dungan, 2012), for recent reviews). For example, the right DLPFC responds when individuals make unfair offers in the ultimatum game (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003) and when allocating money to others (Chang, Smith, Dufwenberg, & Sanfey, 2011; van den Bos, van Dijk, Westenberg, Rombouts, & Crone, 2009) during the trust game. One advantage of studying the DLPFC in social decision-making behaviour is that this region can be targeted by brain stimulation techniques such as transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tdcs). TMS studies of DLPFC stimulation have consistently supported a crucial role for this region in social decision-making. It has been suggested that in these studies the DLPFC serves to down-regulate selfish preferences (Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006).

To gain a fuller understanding of the mechanisms that may support social decisions and link neural responses and behaviour, studies have begun to apply mathematical models borrowed from research in the field of reward-guided behaviour (see **Figure 1.8**). This is a relatively new approach to study social decision-making, which has previously been investigated mainly using economic games such as the prisoners dilemma, trust game and ultimatum game. In economic paradigms participants often experience mixed payoffs and have contrasting goals between self and other. Such a set up can make it difficult to separate self and social preferences, which may be more clearly delineated in reward models. These models of reward-guided decision-making also have a number of advantages when applied to fMRI. Firstly, these models have different internal parameters that can be associated with different precise computations within the same behavioural task. Secondly, they allow us to examine neural responses parametrically rather than relying on traditional subtraction based designs, which average over a number of trials. Third, by relating different parameters together in formal equations, these models can also predict precisely the effect that differences in neural activity should cause in behaviour, reducing the reliance on reverse inference. Finally, hypotheses can be formulated not only about brain regions, but also about the computational mechanisms they may underpin (Behrens et al., 2009).

One of the most replicated findings that have come from mathematical models of neural regions is that midbrain dopamine neurons, which project to the ventral striatum, encode the difference between expected and actual rewards received, rather than just signalling reward outcomes themselves (Schultz, 2013). Such reward prediction error (PE) signalling is the key computation that drives reinforcement learning (RL) models of behaviour (Sutton & Barto, 1998). At their most simple, RL algorithms state that expectations of future reward (Qt+1) should be a function of current expectations (Qt) and their discrepancy from the actual outcome that is experienced—the prediction error (t). These reward predictions are updated by the learning rate (α):

(1) RL model

$$Q_{t(n+1)} = Q_{t(n)} + \alpha \, x \, \delta$$

Where:

(2) *Prediction error* $\delta = r_t - Q_{t(n)}$

In these equations *n* is the number of times an action, *t*, has been performed and *α* is the learning rate, in other words the extent to which the values are updated by new information. In (1) the value of the action in the future ($Q_{t(n+1)}$) is a function of current predicted value of the action (Q_t) added to the prediction error (δ), which is multiplied by the learning rate (α). The learning rate defines the extent to which the prediction error updates the predicted value. Consequently, a low learning rate will minimise the influence of the prediction error and the amount that the value is updated. The prediction error, shown in (2), compares the actual outcome achieved by an action (r) to the prediction of its value ($Q_{t(n)}$). This difference is what determines the updating of the predicted value in the future.

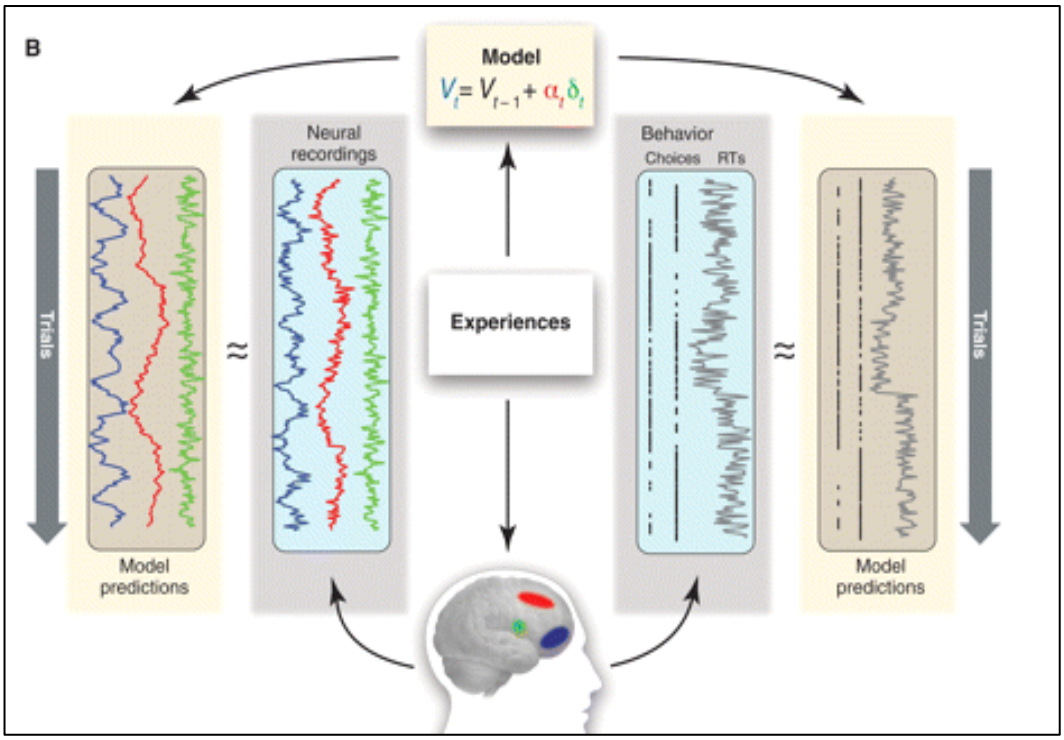
To fit this model to individual participant's behaviour an extra free parameter is needed, the inverse temperature/exploration-exploitation parameter, (β), which is implemented in the softmax function (see 3) and fitted using maximum likelihood estimation. The temperature parameter quantifies the noisiness of participants choice behaviour, with a high parameter meaning very variable choice behaviour and a low parameter meaning very consistent choice behaviour.

(3) Softmax function

$$Pt(r) \quad = \exp\left(\frac{qt(r)}{\beta}\right) / \sum_{n} i = 1 \exp\left(\frac{qt(i)}{\beta}\right)$$

An important aspect of the R-W algorithm and other models based around the principles of RLT, are the assumptions that these models make about learning. By including a free parameter, the learning rate, the extent to which an individual updates the values of a performed action can be defined regardless of the magnitude of the prediction error. Thus, RLT assumes learning is idiosyncratic, allowing for individual differences in learning and decision-making to be accounted for. In addition, this framework also makes the assumption that learning the value of actions is underpinned by the same computational mechanisms, across individuals and across species, as well as in both behaviour and the brain.

Figure 1.8. Schematic of an approach that combines mathematical models of behaviour with neural recordings.



Notes: The model contains parameters that represent specific computations underlying behaviour. As the subject/model undergoes different experiences, these parameters will fluctuate. The fluctuation in these parameters is used to find neural correlates of the specific underlying computations. Separately, the same parameter fluctuations come together to predict changes in behaviour. Reproduced from Behrens et al., (2009) with permission of the copyright owner.

A number of studies have begun to apply the framework of reinforcement learning theory to neural processing in social contexts. In Behrens et al., (2008) participants learned the probability of receiving a rewarding outcome from two options associated with different reward levels. On each trial, participants received advice from a confederate about which option to choose. To maximise financial return participants had to track how volatile the environment was (how rapidly the better option was shifting between the two) and also the volatility of the confederate advice. Whilst response of the ACCs covaried with the environmental volatility, response of the ACCg covaried with the volatility of the social advice at the time of the outcome. These authors also observed responses in the DMPFC, right TPJ/STS and MTG at the time of the outcome and in the sgACC and VMPFC at the time of the decision (Behrens et al., 2008).

In another study prediction errors were created by unexpectedly changing the outcome to be either for the participant themselves or for a charity. This showed that there was PEs in the ventral striatum whether the outcome was for the participant themselves and when the outcome was for a charity (Harbaugh, Mayr, & Burghart, 2007). A separate study observed response in the dorsal striatum to PEs both when participants made a choice and when they observed another person making a choice (Cooper, Dunne, Furey, & O'Doherty, 2012). These studies support the common currency argument that self and other value representations are overlapping. However, one limitation of this common currency interpretation is none of these studies had a 'non-self' control condition. Does the ventral striatum signal a common currency of reward, a general reward prediction error mechanism regardless of whether the agent is self, other or no one? Or a fictive reward signal that reflects a reward that was not delivered to self? (E.g. Hayden, Pearson, & Platt, 2009; Lohrenz, McCabe, Camerer, & Montague, 2007) (This outstanding question will be addressed in **Chapter 6**).

Other studies have suggested that vicarious PEs rely on distinct neurocircuitry. Burke and colleagues used a task that examines self reinforcement-learning and observational reinforcement-learning. Participants performed a probabilistic reinforcement-learning task where they were required to pick one of two abstract stimuli that were probabilistically associated with monetary outcomes. On some trials participants performed individual learning, on some trials they were able to observe the action of a social confederate and in a third condition they observed both the action and outcome of the social confederate. Behaviourally, Burke et al observed that participants were significantly more likely to choose the correct stimuli in the observational learning condition. Activity in the DLPFC correlated with observational action prediction errors whilst activity in the vmPFC corresponded to the observational outcome prediction errors (Burke et al., 2010).

Together, these studies point to a key role for the ventral striatum as well as regions of the prefrontal cortex for guiding social decision-making. Surprisingly, however, no current study has examined neural regions that signal PEs when outcomes are for another person. Moreover, to my knowledge, only a limited

number of studies of social decision-making in humans have used a non-social control condition in order to directly assess the socially-specific vs. common currency conceptions of value-based decision-making. Another issue for the studies that support the shared neural circuits or common-currency theories of vicarious perception and decision-making is that neural overlap does not necessarily mean that the same neurons are responding, or that they are performing the same computations. A recent study examining the common neural representation theory of physical pain and social rejection showed that multivariate pattern analysis was able to separately classify physical pain and social rejection patterns of neural response, and showed that they relied on distinct uncorrelated fMRI patterns within 'pain-processing' regions (Woo et al., 2014). Relatedly, we cannot tell directly from BOLD response the cognitive processes that are being instantiated (Iannetti & Mouraux, 2011). It is often assumed in studies of empathy that 'shared activation' means the participant is experiencing empathy. Finally, social decision-making studies have rarely directly examined trait individual differences to explain variability in social decision-making behaviour. Does empathy predict individual differences in social decision-making behaviour and its neural substrate?

1.6. Outstanding research questions

Empathy is a key component of many aspects of social cognition and behaviour. There are, however, a number of gaps in the extant literature on empathy that have yet to be fully addressed. This thesis identifies four outstanding questions regarding the behavioural and neural foundations of empathic/vicarious processing and investigates these questions empirically.

1.6.1. *How separable are different aspects of empathic processing and do distinct aspects of empathic processing related to different aspects of social functioning?*

As highlighted above, a distinction is often made between 'affective' empathic processes e.g. being aware of, and resonating with, the feelings of another individual such that the awareness of their emotion drives the same state in oneself and 'cognitive' empathic processes e.g. identifying and understanding what another individual is thinking/feeling without a necessary affective response. However, little is known about the degree to which these different aspects of empathic processing relate to one another and whether they are associated with distinct aspects of social functioning in the general population. This lack of knowledge has been compounded by previous studies that have used tasks that do not clearly separate demands for affective and cognitive content. It has also been suggested that individuals with psychopathy and autism may display dissociable empathy impairments (Blair, 2005), but relatively few studies have directly tested this hypothesis and no study to date has related psychopathic and autistic traits in the general population to different aspects of empathic processing. By studying traits associated with these disorders we can potentially 'pull apart' the different components of empathy. The aim of **Chapter 2** is thus to characterise the degree to which different aspects of empathic processing relate to each other and individual differences in psychopathic and autistic traits. An additional aim is to identify whether any of these associations can be explained by variance in alexithymic traits. To this end, I use two established behavioural paradigms measuring affective and cognitive components of empathy (ToM animations and SAM-Faces task) as well as self-report measures of psychopathic, autistic and alexithymic traits.

1.6.2. *How does empathy relate to trait prosocial behaviour and do additional trait constructs moderate the association between empathy and prosocial behaviour?*

Chapter 3 examines (i) associations between different aspects of empathy and prosocial behaviour and (ii) how emotion regulation moderates the association between empathy and prosocial behaviour. It is widely assumed that empathy is an important motivating factor for prosocial behaviour. However surprisingly few empirical studies have addressed the associations between these constructs directly and those that have done so suggest that the association may not be entirely straightforward (Eisenberg, 2000; Singer & Klimecki, 2014). It is also unclear how different components of empathy might relate to prosocial behaviour: is it the capacity to vicariously experience another's' emotion or the capacity to understand another's' thoughts, or both, that motivates prosocial behaviour? Finally, even if empathy does motivate prosocial behaviour, it is unknown whether additional variables may moderate the strength with which empathy and prosocial behaviour are associated. Consequently, in **Chapter 3** I collect trait measures of empathy, prosocial behaviour and emotion regulation to examine associations between these different constructs in a sample of healthy adults (n=110).

1.6.3. *Where in the brain is vicarious information (both negative and positive) processed, and does this vary in:*

1.6.3.1. Children with conduct problems

As highlighted previously, specific neural circuits are thought to be involved in processing information about other people's experiences, such as predictions

and outcomes of pain and reward for others. However, we currently have a limited understanding of how these regions respond in those with disorders of empathy, such as children with conduct problems. It has been suggested that part of the reason these children display antisocial behaviour is due to atypical responses to others' suffering. Yet it is unknown whether neurobiologically children with conduct problems show atypical responses to the observation of pain in others. In **Chapter 4** I examine the neural basis of vicarious processing of other people's pain in a group of children with conduct problems (n=40) compared to a group of typically developing children (n=18) matched for age, IQ and SES, using block-design fMRI. By collecting parent and teacher reports of callous, unemotional and uncaring traits, using the well-validated Inventory of Callous and Unemotional traits (ICU). I also examine whether callousness (which in particular denotes lack of empathy) accounts for the degree of neural response to others' pain in children with conduct problems.

1.6.3.2. Individual differences in typical empathy

The majority of studies measuring the neural basis of vicarious processing have focused on responses to other people's negative experiences, such as pain. There is a paucity of data as to the neural regions involved in processing others' positive experiences, such as predicting and obtaining rewards. A large body of research has examined the neural basis of predictions about the likelihood of receiving rewards ourselves (Schultz, 2013). However, we do not operate in a social vacuum. Successfully cooperating, competing or empathising with others hinges on our ability to predict when others will receive rewards. Yet, very little is known about the neural mechanisms that underpin *vicarious* reward prediction. In addition, no existing studies have measured how individual differences in empathy might explain variability in response to other people's rewards. We know that these are large individual differences in the level of empathy between individuals. Do these individual differences also modulate how social information is processed in the brain? In **Chapter 5** I use event-

related fMRI in 30 healthy male volunteers (age 18-32) to investigate neural responses as participants view cues that indicate either a high or low probability of reward for themselves or for a social confederate. I also collect trait measures of empathy (QCAE) to examine whether this measure can explain variability in response of neural regions that process vicarious reward prediction.

1.6.4. *What regions of the brain are involved in signaling prosocial prediction errors? What are the mechanisms that link empathy to prosocial decision-making behaviour and neural response?*

In the final empirical chapter, **Chapter 6**, I examine the behavioural and neural mechanisms that might link individual differences in empathy to variability in prosocial decision-making behaviour. As mentioned previously, reinforcement learning theory has provided a detailed characterisation of neural computations that underlie reward-guided behaviour. Specifically, in reinforcement learning theory, prediction errors – the difference between a predicted and actual outcome of a choice – act as a neural signal to drive learning. The neural and behavioural mechanisms of learning from reinforcement delivered to oneself are increasingly well understood. However, less is known about how we process rewards delivered to other people. To be motivated to make beneficial decisions for another person (i.e. behave prosocially), it may be critical to represent and vicariously process rewards that others receive following our choices. Despite this, to date no studies have, to my knowledge, examined the neural mechanisms that underlie learning to obtain rewarding outcomes for others and how this may vary with trait empathy.

Thus in **Chapter 6** thirty-one healthy male participants (aged 18-32) performed a probabilistic reinforcement learning task in which they were required to learn the probability that each of two stimuli (high probability vs. low probability) would be rewarded. Participants performed this task for themselves (self reinforcement condition), for another participant (confederate, prosocial learning

condition) or for no one (no reinforcement, control condition). Participants also completed a questionnaire measure of trait empathy (QCAE). Participants' behaviour was modelled using a Rescorla-Wagner (R-W) computational model. The model assumes that on each trial, participants make a prediction of the value of a stimulus, which is updated by a prediction error – the difference between the value of the reward they expected and the value of the reward they received. The amount that the prediction is updated is dictated by the learning rate, which quantifies how quickly participants learn from experience and can be different across different individuals and different contexts. Behaviourally, I examine differences in learning rates in the three learning conditions (self, prosocial and no one) as well as whether empathy modulates the prosocial-self difference in these parameters. At the time of the outcome, the PE parameter is used as a parametric modulator to examine common and distinct neural processing of self, prosocial and no one PEs in neural regions previously implicated in social decisions/social PEs including the ventral striatum, sgACC, DLPFC, OFC and ACCg. Finally, individual variability in PE signalling at the time of the outcome is examined in relation to trait empathy.

CHAPTER 2: Chapter background

As set out in the introduction of this thesis, empathy is a multidimensional phenomenon comprised of different social cognitive processes. A distinction is often drawn between 'affective' empathic processes e.g. being aware of, and resonating with, the feelings of another individual such that the awareness of their emotion drives the same state in oneself (henceforth affective resonance in this chapter) and 'cognitive' empathic processes e.g. identifying and understanding what another individual is thinking/feeling without a necessary affective response (henceforth cognitive perspective-taking in this chapter). However, little is known about the degree to which these different components are associated with one another and whether they differentially relate to psychopathic and autistic traits in the general population. In the current study I describe two behavioral paradigms and measures of individual differences in social-cognitive functioning (psychopathic, autistic and alexithymic traits) that were employed to address these outstanding research questions.

2.1. Introduction

2.1.1. *Psychopathy as a disorder of empathy*

As described in the introduction, psychopathy is a disorder characterised by a lack of empathy, shallow affect, and manipulation of others for own gain (Hare, 1999). In individuals with psychopathy, difficulties with affective resonance are often apparent. For example, individuals with psychopathy show reduced physiological response to others' distress (Blair et al.,1997). Adults with psychopathy display atypical neural responses to others' pain (e.g. Decety et al., 2013). In community samples, high levels of psychopathic traits are related to weaker affective responses to fearful faces and happy stories (Seara Cardoso et al., 2012; Seara Cardoso et al., 2013) as well as being negatively associated with the enjoyment of prosocial interactions (Foulkes et al., 2014). Taken together, these findings indicate clear difficulties in resonating with others' emotions in both clinical samples with psychopathy and in community individuals with high levels of psychopathic traits.

In contrast, one of the defining features of psychopathy is the ability to successfully manipulate others (Hare, 2003). Thus it might be expected that psychopathy would be associated with typical cognitive perspective-taking. Several studies report no cognitive perspective-taking impairments (Anastassiou-Hadjicharalambous & Warden, 2008; Blair et al., 1996; Dolan & Fullam, 2004; Richell et al., 2003) and even superior ability (Hansen et al., 2008) in individuals with psychopathy or high psychopathic traits. However, others have reported problems with tasks related to cognitive perspective-taking in both incarcerated psychopaths (Brook & Kosson, 2013) and healthy samples with high psychopathic traits (Ali & Chamorro-Premuzic, 2010). One possibility for these mixed findings is that different paradigms vary in their level of affective content, with some purported cognitive perspective-taking measures requiring identification of other people's feelings, rather than just their thoughts. It could

be that negative associations between psychopathic traits and cognitive perspective-taking are driven by problems related to basic affective processing, rather than difficulties in cognitive perspective-taking per se. In fact, all studies that have reported that psychopathy/psychopathic traits are associated with poorer cognitive perspective-taking have utilised measures with affective content (e.g. Ali & Chamorro-Premuzic, 2010; Brook & Kosson, 2013) and therefore do not necessarily provide evidence for cognitive perspective-taking impairments in psychopathy.

2.1.2. *Autism as a disorder of empathy*

Autism spectrum disorders (henceforth ASD) are characterised by problems with social interaction, communication and repetitive behaviours. It has been argued that individuals with ASD also have reduced empathy (e.g.(Baron-Cohen & Wheelwright, 2004). Several decades of research indicate that individuals with ASD have difficulties with cognitive perspective-taking in particular are central to ASD (see Hill & Frith, 2003). The findings from studies assessing processes related to affective resonance in ASD are less consistent. There is evidence of absent sensori-motor resonance when viewing others' pain in individuals with ASD (Minio-Paluello et al., 2009). However, other studies have shown typical sensori-motor resonance when viewing others in pain (Fan et al., 2013) and appropriate physiological responses to others distress (Blair, 1999) in individuals with ASD. When cognitive perspective-taking and empathic concern, a process related to affective resonance, have been compared in individuals with ASD, impairments in cognitive perspective-taking but not empathic concern were found (Dziobek et al., 2008). Some theorists have argued that affective resonance is actually heightened in individuals with ASD (Smith, 2009) and reports of greater empathic facial affect in children with ASD compared to controls supports this (Capps et al., 1993).

2.1.3. *Alexithymia as a disorder of empathy*

A further consideration for the profile of empathy in ASD is the high comorbidity of the disorder with alexithymia. Alexithymia is a sub-clinical condition defined by an inability to identify and describe feelings in the self. Preliminary behavioural and neuroimaging research suggests that affective and empathy impairments in ASD may be a function of interoceptive difficulties related to alexithymia rather than ASD per se (Bird et al., 2010; Silani et al., 2008) and that after accounting for the variance explained by alexithymia in these studies of ASD there is no difference in empathy between individuals with ASD and controls (Bird & Cook, 2013). However, one recent fMRI study found no significant moderating effects of alexithymia in a task where participants viewed pictures of other people in pain (termed an 'empathy for pain' task by the authors) in individuals with ASD (Fan et al., 2013). Nevertheless, the variance in alexithymia scores was very limited (SD 3.8 in Fan et al., 2013 vs 11.8 in Bird et al., 2010), which may explain why no effect of alexithymia was observed. Less is known about the possible contribution of alexithymia to empathy impairments seen in individuals with psychopathy. Although the co-occurrence rates of alexithymia and psychopathy are lower than for ASD (Louth, Hare, & Linden, 1998), the two disorders do share some common attributes (Lander, Lutz-Zois, Rye, & Goodnight, 2012).

2.1.4. *Associations between psychopathy, autism and alexithymia*

To date, only two studies have directly compared the profile of affective and cognitive processing related to psychopathy and ASD, and these have both been in children. Children with conduct disorder and psychopathic traits showed less affective resonance with others' emotions but did not have problems with cognitive perspective-taking; conversely, children with ASD showed reduced cognitive perspective-taking but did not have problems with affective resonance

(Jones et al., 2010; Schwenck et al., 2012). However, no studies have directly contrasted psychopathic and ASD traits and processes related to affective resonance and cognitive perspective-taking in adults. Moreover, no studies have investigated the contribution of alexithymia to ASD and psychopathic traits in tandem. Psychopathic, ASD and alexithymic traits are present in varying degrees in the general population (Bagby, Parker, & Taylor, 1994; Baron-Cohen et al., 2001; Hare & Neumann, 2008). Taxometric studies indicate that psychopathy should be viewed as a dimensional construct that is an extreme variant of normal personality and not a distinct category of behavior (see Hare & Newman, 2008 for review). Similarly, behavioral genetic studies indicate a similar etiology of autistic traits in the general population as well as in clinical groups (Robinson et al., 2011), thus providing an empirical basis for studying variants in traits associated with these disorders in the general population. Finally, investigating associations between these traits and potential differences in social information processing is one way to dissect the component processes that may contribute to empathy, a key aim of the present chapter.

2.1.5. *The current study*

The present study investigated (i) whether psychopathic and ASD traits were differentially related to performance on affective resonance and cognitive perspective-taking tasks and (ii) whether alexithymia contributes to task performance. Based on previous research, I predicted that psychopathic traits would be negatively associated with performance on the affective resonance task but not the cognitive perspective-taking task and that ASD traits would be negatively associated with performance on the cognitive perspective-taking task but not the affective resonance task. Alexithymia has previously been demonstrated to predict empathy deficits while recent neuroimaging results suggest cognitive perspective-taking is unlikely to be affected (Bernhardt et al., 2014). Therefore, I predicted that alexithymia would make a contribution to performance on the affective resonance task, but be unrelated to performance

on the cognitive perspective-taking task.

2.2. Materials and Methods

2.2.1. *Participants*

One-hundred-and-ten healthy adults (50% M; 50% F) aged 18-33 (*M* = 21.9, *SD* = 3.7) with estimated IQ between 87-129 (M=116.8, SD=8.4) took part. Participants were recruited through university participant databases and the community. All participants provided written informed consent and the study had institutional ethics approval.

2.2.2. *Procedure*

Participants completed two tasks to assess affective resonance and cognitive perspective-taking as part of a larger battery of tasks. All tasks were presented in a randomised order followed by the questionnaires.

2.2.3. *Experimental tasks*

2.2.3.1. Theory of mind animations task (Cognitive perspective-taking task)

This task assessed participants' ability to understand others' complex mental states (e.g. tricking, coaxing) and has been previously used to examine ToM abilities in children with autism (Abell, Happé and Firth, 2000) and healthy participants (Castelli, Frith, Happé & Frith, 2002). I selected four 'ToM' and four 'random' animations from Abell et al (2000). Each animation featured two characters: a big red and small blue triangle either interacting with one another (ToM animations) or moving randomly (random animations). Participants were asked to watch each animation carefully and to describe what was happening whilst their verbal responses were recorded. Two people transcribed the verbal descriptions that were coded in terms of intentionality and appropriateness. The intentionality scale ranged from 0 (no appreciation of another agent, nor actions or mental states) to 5 (the agent acts with the goal of affecting or manipulating the other agent's mental states). The appropriateness scale ranged from 0-3. One researcher rated all transcriptions and a second researcher rated a random sample of 56. Intra-class correlations (ICC) between raters for intentionality (ICC, single measures = .682) and appropriateness (ICC single measures = .760) were good. The ratings of intentionality and appropriateness were converted to z scores and a composite score was created.

2.2.3.2. Self-assessment manikin faces task (Affective resonance task)

This task assessed participants' affective empathic response to emotional faces using the SAM rating scale (Seara-Cardoso et al., 2012). Participants were required to rate their own emotional response to the affective state of another on a 9-point manikin (changing from smiling to a sad face with a neutral expression in the middle) whilst viewing images depicting a person showing either a sad, fearful, angry, happy or neutral expression. The order of images was randomised for each participant. Ratings for sad, fear and anger were reverse scored so that the higher scores reflected ratings of greater distress, and thus greater affective resonance, when viewing others' negative emotions.

These variables were then converted to z-scores and a composite score was created along with happy ratings.

2.2.4. *Questionnaires*

2.2.4.1. Assessment of psychopathic traits

Psychopathic traits were assessed with the SRP-4-SF short form (SRP-4-SF; Paulhus et al., 2015), a 29-item scale designed to measure psychopathic attributes in non-institutionalised samples. The SRP has been shown to have good construct validity and internal consistency (Cronbach's alpha .89 in the present study) and is strongly correlated with the PCL-R; the clinical measure of psychopathy (Lilienfeld, Fowler, & Patrick, 2006; Paulhus et al., 2015). Questions were rated on a 5-point scale from "Disagree Strongly" to "Agree Strongly" and included items such as "Most people are wimps" and "I love violent sports and movies".

2.2.4.2. Assessment of autistic traits

ASD traits were assessed with the Autism Quotient (AQ; Baron-Cohen et al., 2001), a 50 item scale designed to assess ASD traits in both clinical and community samples. The AQ has good construct validity and internal consistency (Cronbach's alpha .83 in the present study). Questions were rated on a 4-point scale from "Definitely Disagree" to "Definitely Agree" and included items such as "I enjoy meeting new people" and "I would rather go to a library than a party".

2.2.4.3. Assessment of alexithymia

Alexithymic traits were assessed with the Toronto Alexithymia Scale (TAS, (Bagby et al., 1994) a 20-item scale designed to measure subclinical alexithymic traits. Questions were rated on a 5-point scale from “I Strongly Disagree” to “I Strongly Agree” and included items such as “I am often confused about what emotion I am feeling” and “I am often puzzled by sensations in my body”. The TAS has good construct validity and internal consistency (Cronbach’s alpha .82 in the present study).

2.3. Results

2.3.1. *Correlations between measures*

Performance on the affective resonance and cognitive perspective-taking tasks was positively correlated (r = .40, p < .001). All questionnaire measures were also positively correlated with one another (See **Table 2.1**). First, bivariate correlations were examined to assess whether psychopathic and ASD traits were differentially related to affective resonance and cognitive perspective-taking. As predicted psychopathic traits showed a statistically significant negative correlation with performance on the affective resonance task (r = -.258, p = .02) whilst ASD traits did not (r = -.102, p = .291). Conversely, ASD traits showed a statistically significant negative correlation with performance on the cognitive perspective-taking task (r = -.209, p = .028) whilst psychopathic traits did not (r = -.046, p = .634).

Table 2.1 Correlations between questionnaire measures of psychopathic, autism spectrum disorder and alexithymic traits and task performance.

	SRP	AQ	TAS	AR
AQ	.244*			
TAS	.252*	.370**		
AR	-.258**	-.102	-.245*	
CPT	-.046	-.209*	-.120	.399**

Abbreviations: SRP, Self-Report Psychopathy Scale; TAS, Toronto Alexithymia Scale; AQ, Autism Quotient; AE, Affective Resonance task; CPT, Cognitive Perspective-Taking task

* p<.05

** p<.01

2.3.2. *Regression analyses*

I conducted hierarchical multiple regression analyses to investigate whether psychopathic and ASD traits were uniquely and differentially related to affective resonance and cognitive perspective-taking, and to examine whether individual differences in alexithymia and/or IQ might explain any associations (see **Table 2.2**).

Table 2.2. Hierarchical multiple regression between questionnaire measures of psychopathic, autism spectrum disorder and alexithymic traits and task performance.

	Affective Resonance task				Cognitive Perspective-Taking task		
	Beta	t	p		Beta	t	p
Step 1				Step 1			
SRP	-.258	-2.772	.007*	AQ	-.209	-2.224	.028*
Step 2				Step 2			
SRP	-.248	-2.574	.011*	AQ	-.211	-2.16	.033*
AQ	-.041	-.428	.669	SRP	.005	.056	.956
Step 3				Step 3			
SRP	-.213	-2.209	.029*	AQ	-.193	-1.868	.065^
AQ	.025	.245	.807	SRP	.014	.144	0.885
TAS	-.201	-1.991	.049*	TAS	-.052	-.501	0.618
Step 4				Step 4			
SRP	-.218	-2.227	.028*	AQ	-.196	-1.895	.061^
AQ	.024	.236	.814	SRP	-.000	.000	1.000
TAS	-.200	-1.977	.051^	TAS	-.050	-.483	.630
IQ	.033	.353	.725	IQ	.106	1.113	.268

^ p<.10, * p<.05

Abbreviations: SRP, Self-Report Psychopathy Scale; TAS, Toronto Alexithymia Scale; AQ, Autism Quotient ;Full IQ calculated from Weschler Intelligence Test of Adult Reading

Two models were run. For the model predicting performance on the affective resonance task, psychopathic traits were entered at the first stage. Psychopathic traits significantly predicted reduced affective resonance (p = .007). At the second stage ASD traits were entered. Psychopathic traits were uniquely negatively associated with affective resonance (t = -2.57, p = .011) whilst ASD traits were not (t = -.43 p = .669). The R-squared change was not significant (F change = .18, p = .669) indicating that ASD traits did not explain significantly more variance in the model. At the third stage, alexithymia scores were entered. Controlling for alexithymia did not change the pattern of results, but there was a unique negative association between alexithymia and affective resonance (t = -1.99, p = .049), and the R-squared change was significant (F = 3.96, p = .049). At the fourth stage IQ scores were entered. Controlling for IQ did not change the pattern of results, nor was IQ a significant predictor of affective resonance (p = .73). The same regression sequence was then used for cognitive perspective-taking, but with ASD traits at the first stage and psychopathic traits at the second. ASD traits were significantly negatively associated with cognitive perspective-taking (t = -2.22, p = .028). At the second stage psychopathic traits were entered. ASD traits were uniquely negatively associated with reduced cognitive perspective taking (t = - 2.16, p = .033) whilst

psychopathic traits were not (t = .06, p = .956). The R-squared change was not significant (F change = .00, p = .956) indicating that psychopathic traits did not explain significantly more variance in the model. Taking into account alexithymia and IQ did not change the pattern of results, nor did either of these variables predict cognitive perspective-taking. No further R-squared changes were significant (all F's < 1.24, all *ps* > .26).

2.4. Discussion

The current chapter compared associations between psychopathic and ASD traits and tasks assessing affective resonance or cognitive perspective-taking. I demonstrated unique associations between psychopathic traits and reduced affective resonance but not cognitive perspective-taking, and unique associations between ASD traits and reduced cognitive perspective-taking but not affective resonance. Alexithymic traits did not explain observed associations between task performance and psychopathic or ASD traits but rather contributed to performance on the affective resonance task independently of psychopathic traits. This is the first study in healthy adults to show a differential relationship between these variables. Thus, it extends previous findings that have reported contrasting profiles of empathy impairments between children with psychopathic tendencies or ASD (Jones et al., 2010; Schwenck et al., 2012). These results also suggest that although affective resonance and cognitive perspective-taking measures share some variance, they can capture dissociable processes, and thus extends our knowledge regarding the structure of empathy.

Psychopathy is thought to be characterised by problems with affective resonance but not cognitive perspective-taking. I used measures that were designed to specifically probe affective resonance and cognitive perspective-taking, without there being cognitive perspective-taking demands on the affective resonance task or vice versa. These results therefore extend and clarify the findings of previous studies reporting reduced affective resonance in

individuals high in psychopathic traits (Seara Cardoso et al., 2012; Seara Cardoso et al., 2013) by indicating a reduction in affective resonance in the absence of a reduction in cognitive perspective-taking. These data also highlight how high psychopathic traits are not related to atypical cognitive perspective-taking when a task without an affective component is used.

ASDs have been consistently linked to problems with cognitive perspective-taking (Hill & Frith, 2003). Interestingly, I found that elevated ASD traits in the general population were also associated with atypical cognitive perspective-taking. In contrast, findings of tasks related to affective resonance processing in ASD are mixed, with reduced (Minio-Paluello et al., 2009), intact (Blair, 1999; Dziobek et al., 2008; Bird et al., 2010; Fan et al., 2013) and elevated (Capps et al., 1993) levels of affective processing being reported. These findings suggest that ASD traits are not associated with either a reduced or an enhanced ability to resonate with the emotions of another, despite the fact that high levels of ASD traits are related to difficulties with understanding others' minds. It would be useful for future studies to assess multiple forms of processing related to affective resonance, as the paradigms used in some studies that reported intact affective resonance investigated empathic concern, rather than affective resonance. Examining both of these processes in tandem may help to shed further light on the profile of empathic processing in ASD. Moreover, it would also be of interest to further examine the exact cognitive perspective-taking mechanisms that may be disrupted in relation to ASD/high ASD traits. It could be that some disrupted components of cognitive perspective-taking relate to bottom-up processes such as detection of biological movement, whereas others might relate to top-down processes such as the influence of situational cues.

Both psychopathy and ASD have previously been associated with elevated levels of alexithymia (Bird & Cook, 2013; Louth et al., 1998; Lander et al., 2012), and I also observed modest correlations between psychopathic and ASD traits with alexithymia in the present study. Nevertheless, controlling for alexithymic traits did not change the reported associations between psychopathic traits and reduced affective resonance or ASD traits and reduced cognitive perspective-taking. In other words, the reduced ability to identify and describe feelings in the self did not account for the relationship between

psychopathic traits and affective resonance or ASD traits and cognitive perspective-taking. The finding that alexithymia did not explain the reduced cognitive perspective-taking abilities characteristic of ASD traits is of particular interest given recent evidence and theory suggesting that alexithymia does account for affective processing deficits related to autism, when they are observed (Bird & Cook, 2013). These data extend this account by showing that alexithymia does not appear to explain reduced cognitive perspective-taking related to high ASD traits.

I also found that alexithymic traits were negatively associated with a reduction in affective resonance independently of psychopathic traits. This suggests that reductions in affective resonance can be affected both by reduced ability to identify and describe feelings (a characteristic of alexithymia) and a reduced tendency to feel what others feel (a characteristic of psychopathy) and consequently, that there could be different component processes within the construct of affective resonance. Future studies could help to determine the mechanisms underlying reduced affective resonance in psychopathy and alexithymia.

2.4.1. *Limitations and future directions*

A few limitations to the present study should be highlighted. In everyday life empathic responses to others occur in the context of social interactions, the present tasks did not present such scenarios in the interest of isolating affective resonance and cognitive perspective-taking demands. However, this will be explored further in **Chapters 5 and 6.** Although I chose paradigms to specifically examine two process that contribute to the experience of empathy, these are not exhaustive and further research would benefit from examining a larger collection of tasks that tap a multitude of processes related to empathy i.e. empathic concern, mimicry and identifying others emotions. It will also be of interest to determine whether the processing atypicalities associated with psychopathic, ASD and alexithymia traits explain real life observations of

unempathic behavior, as rated by others or observed in an experimental setting. Finally, replication of these results with clinical populations would be informative.

2.4.2. *Conclusions*

This chapter described a behavioural study where I investigated how different aspects of empathic processing can be associated and dissociated in the general population. Overall, the findings are that two key components of empathic processing: affective resonance and cognitive perspective-taking, show a modest positive association in the general population. However, clear distinctions are also evident. This study showed, for the first time, that elevated psychopathic traits are related to reduced affective resonance but not cognitive perspective-taking, whilst elevated levels of ASD traits are related to reduced cognitive perspective-taking but not affective resonance in a community sample of adults. Consequently, these data suggest that although some level of 'empathic competence' appears to generalise across paradigms, it also appears that different social information processes (potentially underpinned by separable neural circuits) account for individual differences in different types of empathic/social behaviour.

In the next chapter, **Chapter 3**, I will turn to the question of the degree to which empathy predicts prosocial behavior. I will also examine whether additional variables, namely the capacity to regulate one's own emotional experience, moderates the strength with which empathy is predictive of prosocial behaviour.

CHAPTER 3: Chapter background

Theory and evidence suggest that empathy is an important motivating factor for prosocial behaviour, i.e. social behaviour intended to benefit another. However, despite the general assumption that empathy will motivate prosocial behaviour, surprisingly few studies have addressed this question directly. It is also unclear how different components of empathy might relate to prosocial behaviour: is it the capacity to vicariously experience another's' emotion or the capacity to understand another's' thoughts, or both, that motivates prosocial behaviour? Finally, even if empathy does motivate prosocial behaviour, it is unknown whether additional variables may moderate the strength with which empathy and prosocial behaviour are associated. The current study set out to address these outstanding questions. A sample of healthy adults (N=110) completed questionnaire measures of empathy and prosocial tendencies, as well as a questionnaire measure of their ability to regulate their own emotions. I examined how different components of empathy and prosocial tendencies were related, as well as whether emotion regulation moderated the degree to which these constructs were associated. Whilst in the previous chapter (**Chapter 2**) I interrogated two specific aspects of empathy, affective resonance and cognitive perspective taking, using behavioural tasks, in the current chapter I used trait measures of empathy that probe self-reports of behaviours, feelings and reflections that give a proxy for affective and cognitive empathy computations.

3.1. Introduction

3.1.1. *What is prosocial behaviour?*

Prosocial behaviours can be broadly defined as social behaviours intended to benefit other people. Humans have a remarkable capacity to engage in prosocial behaviours, even with genetically unrelated individuals (Fehr & Fischbacher, 2003). For example, people routinely engage in charitable donation behaviour and exhibit social preferences, where their preferred choices are based on a positive or negative concern for the welfare of others (Fehr & Camerer, 2007). However, the processes that influence how and when prosocial behaviours occur remain poorly understood. Theory and evidence have suggested that empathy is one of the key motivating factors for prosocial behaviour (Eisenberg et al., 2010; Hoffman, 2008; Penner et al., 2005), but this assertion has received little empirical attention to date.

3.1.2. *Empathy and prosocial behavior*

There is evidence that processes related to both affective and cognitive empathy are positively associated with prosocial behaviour (for a review see (Eisenberg et al., 2010). The majority of these studies have used the interpersonal reactivity index (IRI; Davis, 1983), which measures trait empathic concern/sympathy, or cardiovascular and electrodermal indices, such as heart rate deceleration and facial electromyographic (EMG), as proxy measures of affective empathy. For example, heart rate deceleration (which is thought to index vicariously induced sadness or sympathy, e.g. (Eisenberg et al., 1988), and increased indicators of facial sadness when watching needy others are associated with increased willingness to help (Eisenberg et al., 1989).

Dispositional empathic concern, as measured by the IRI, has also been linked to higher levels of self-reported charitable giving (Davis, 1983) and greater self-reported concern for the welfare of others (Batson, 1998). In terms of associations between cognitive components of empathy and prosocial behaviour, studies have focused on correlating the perspective-taking subscale of the IRI to self-reported prosocial behaviour and have found that trait perspective taking is positively associated with frequency of volunteering (Carlo, Allen, & Buhman, 1999) and self-reported prosocial tendencies (Carlo, Hausmann, Christiansen, & Randall, 2003). It should be noted, however, that the empathic concern and perspective taking scales of the IRI tap constructs that, although related, are different from the current conceptualisation of 'affective empathy' and 'cognitive empathy' (Singer & Lamm, 2009). Nonetheless, together, these studies broadly suggest that affective and cognitive empathic components may motivate prosocial behaviour.

3.1.3. *The influence of moderating variables*

Whilst it is often assumed that an empathic response to another's distress will motivate prosocial behaviour, Eisenberg (2000) points out that association between the two constructs are often modest and sometimes weak. A possible reason for these modest associations is the influence of moderating variables (Eisenberg, 2000). It has been suggested that emotion regulation, i.e. the capacity to modulate or exert control over an emotional response, might be one such moderator variable (Eisenberg & Fabes, 1992; Hoffman, 2001). Eisenberg and Fabes (1992) propose a model whereby individual differences in both the emotional intensity and regulation capacities are related to an individual's level of prosocial responding. Specifically, they suggest that the perception of distress in another leads to emotional arousal, but emotion regulation i.e. and how this arousal is evaluated by the observer, will influence the subsequent goal directed behaviour, either to improve their own situation or help the others' situation (Eisenberg & Fabes, 1992). The degree of emotion regulation during a

state of emotional arousal (over-, optimal-, or under-regulation) is also proposed to relate to the likelihood of prosocial behaviour. For example, individuals who are able to optimally regulate their arousal, so that they do not experience undue distress in the face of another person's emotions and thus do not become self-focused, are proposed to behave prosocially (Eisenberg & Fabes, 1992). In contrast, individuals who are over- regulated are proposed to exhibit proactive withdrawal, which inhibits prosocial behaviour. Finally, those who are under-regulated are proposed to be prone to aggression and thus more likely to exhibit antisocial rather than prosocial behaviour in an emotionally arousing situation (Eisenberg & Fabes, 1992).

3.1.4. *Different types of emotion regulation*

The model outlined by Eisenberg and Fabes (1992) discusses the degree of emotion regulation (over-, optimal-, or under-regulation) as important for linking empathy to prosocial behaviour. However, it is also likely that the type of emotion regulation strategy used will be important. Both cognitive reappraisal and expressive suppression represent emotion regulation strategies (Gross, 2013; Gross & John, 2003; Gross & Thompson, 2007). Cognitive reappraisal involves reinterpreting an emotional response so that the intensity of its emotional impact is modified (Gross & John, 2003). For example, re-framing a distressing situation as a situation where someone will benefit from support, as opposed to a situation where someone is emotionally labile and potentially unpleasant. Consequently, cognitive reappraisal will enable a person to focus on strategies to provide constructive helping behaviours, rather than the aversive qualities of the situation. Cognitive reappraisal is thought to be a successful emotion regulation strategy, decreasing negative affect and resulting in an attenuation of blood pressure (Ray et al., 2005; Richards & Gross, 2000).

In contrast, expressive suppression involves actively inhibiting on-going emotion-expressive behaviour (Gross, 1998, 2013; Gross & Thompson, 2007). For example, managing an emotional response to an aversive situation in an

effortful manner such that cognitive resources are consumed. Expressive suppression is thought to be a suboptimal strategy because it creates a conflict between heightened emotional arousal and overt expression of the arousal (Gross & John, 2003; Gross & Levenson, 1993; John & Gross, 2004; Soto, Perez, Kim, Lee, & Minnick, 2011). These two types of emotion regulation strategies also appear to lead to different outcomes and consequences for interpersonal functioning (Gross, 2013; Gross & John, 2003; Gross & Levenson, 1993; John & Gross, 2004; Soto et al., 2011). Whilst cognitive reappraisal is positively related to having closer relationships with friends, fewer depressive symptoms and greater life satisfaction, expressive suppression is associated with greater experience of negative emotions, disturbed interpersonal interactions, avoidance of close relationships and reports of less life satisfaction and optimism (Gross, 2013; Gross & John, 2003; Gross & Levenson, 1993; John & Gross, 2004; Soto et al., 2011).

3.1.5. *The current study*

Despite the evidence linking empathy to prosocial behaviour (e.g. Carlo et al., 1999; Eisenberg et al., 1989) and the proposal that individual differences in emotion regulation may moderate associations between empathy and prosocial behaviour (Eisenberg & Fabes, 1992; Hoffman, 2001), this has not, to my knowledge, been directly examined. Moreover, how distinct emotion regulation strategies might moderate associations between empathy and prosocial behaviour has not been explored.

Importantly, the majority of studies suggesting empathy as a motivating factor for prosocial behaviour have investigated self-reported empathic concern (feeling ‘for’ another person, including compassion and sympathy, e.g. Batson, 1998; Davis, 1983), rather than self-reported affective empathic responses (the ability to vicariously experience the emotional experience of others; or feeling ‘as’ another individual). While these two processes are no doubt closely related, there is a lack of empirical data regarding how feeling in a similar emotional

state to another may motivate prosocial behaviour. In addition, self-reported cognitive empathic ability (i.e. the ability to position oneself 'in another person's shoes') might also relate to prosocial behaviour, but compared to the role of affective empathic processes motivating empathy this has received relatively little attention to date (c.f. Carlo et al., 1999; Carlo et al., 2003).

3.1.6. *Hypotheses*

On the basis of previous research and theory (e.g. Batson, 1998; Carlo et al., 2003; Eisenberg et al., 2010), I predicted that both dispositional cognitive and affective components of empathy would be associated with increased prosocial tendencies, but the amount of variance in prosocial behaviour explained by the two types of empathy may be unequal. I also tested interactions between the components of empathy (affective and cognitive) and types of emotion regulation strategy (cognitive reappraisal and expressive suppression) to examine whether individual differences in emotion regulation strategy moderate associations between empathy and prosocial behaviour.

3.2. Materials and Methods

3.2.1. *Participants*

One-hundred-and-ten healthy adults (50% males; 50% females) aged 18-33 (M=21.9, SD=3.7) were recruited through university participant databases (comprised of undergraduate and postgraduate students as well as non-student community members) and through online advertisement. Exclusion criteria included previous or current neurological or psychiatric disorder (as reported by

the participants) and non-normal or non-corrected to normal vision. Participants were compensated at a rate of £8 per hour. All participants provided written informed consent and the study was approved by the University College London Clinical, Educational and Health Psychology Research Ethics committee.

3.2.2. *Procedure*

Participants completed questionnaires to assess empathy, emotion regulation and prosocial tendencies as part of a larger battery of tasks and questionnaires.

3.2.2.1. Assessment of empathy

The Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers, Corcoran, Drake, Shryane, & Vollm, 2011) is a multidimensional empathy questionnaire devised to measure the ability to comprehend the emotions of another (cognitive empathy) as well as the ability to vicariously experience the emotional experience of others (affective empathy). In the development of the QCAE, two raters selected items from other well-validated and commonly used empathy measures (e.g. Empathy Quotient; Baron-Cohen & Wheelwright, 2004), Hogan Empathy Scale (Hogan, 1969); the Empathy subscale of the Impulsiveness-Venturesomeness-Empathy Inventory; (Eysenck & Eysenck, 1978) and the IRI; Davis, 1983) if they were deemed to measure affective or cognitive empathy. Items from these scales deemed to measure other processes (e.g. sympathy) were not included. A Principal Component Analysis of the selected items revealed five sub-scales, further organized in two components assessing cognitive and affective empathy. The cognitive empathy component comprises subscales measuring perspective-taking (e.g. “I am good at predicting how someone will feel”) and Online simulation (e.g. “Before

criticizing somebody, I try to imagine how I would feel if I was in their place."). The affective empathy component assesses emotion contagion (e.g. "People I am with have a strong influence on my mood"); peripheral responsivity (e.g. "I usually stay emotionally detached when watching a film"); and proximal responsivity (e.g. "I often get emotionally involved with my friends' problems"). Items are rated on a 4-point scale from "strongly disagree" to "strongly agree". The QCAE has good validity and internal consistency (Reniers et al., 2011). In the present study Cronbach's alpha for cognitive empathy component .87; affective empathy component .88).

3.2.2.2. Assessment of prosocial tendencies

The Prosocial Tendencies Measure (PTM; Carlo & Randall, 2002) is a 23-item self-report measure that assesses various prosocial tendencies such as compliant prosocial tendencies (e.g. "When people ask me to help them, I don't hesitate"), dire prosocial tendencies (e.g. "I tend to help people who hurt themselves badly") and emotional prosocial tendencies (e.g. "I tend to help others particularly when they are emotionally distressed"). Items are rated on a 5-point scale from "Does not describe me at all" to "Describes me greatly". The PTM has good construct validity and internal consistency (Carlo & Randall, 2002; in the present study Cronbach's alpha .86).

3.2.2.3. Assessment of emotion regulation

The Emotion Regulation Questionnaire (ERQ; Gross & John, 2003) is comprised of two dimensions that assess either reappraisal or suppression regulation strategies. The reappraisal dimension contains items such as "I control my emotions by changing the way I think about the situation I'm in" and the suppression dimension has items such as"I control my emotions by not

expressing them". Items are rated on a 7-point scale from "Strongly disagree" to "Strongly agree". The ERQ has good construct validity and internal consistency (Gross & John, 2003); in the present study Cronbach's alpha for reappraisal subscale .73; suppression subscale .87).

3.2.3. *Statistical analyses*

Bivariate correlations were corrected for multiple comparisons using Benjamini & Hochberg False Discovery Rate (Benjamini & Hochberg, 1995). Corrected p-values are reported. Steiger's *Z* tests (two-tailed) were conducted to test if the different types of empathy (i.e. affective and cognitive empathy) and the different types of emotion regulation strategies (i.e. cognitive reappraisal and expressive suppression) presented differential correlation coefficients with prosocial tendencies. Moderation analyses were then conducted to investigate whether the affective or cognitive empathy subscales interacted with either types of emotion regulation (reappraisal or suppression) to predict prosocial tendencies. All predictor variables were mean centred prior to analyses. Separate regression models using either the affective empathy component of the QCAE (QCAE-affective empathy) or the cognitive empathy component of the QCAE (QCAE-cognitive empathy) at the first stage; the reappraisal subscale of the ERQ (ERQ-reappraisal) or the suppression subscale of the ERQ (ERQ-suppression) at the second stage and the interaction term between these variables at the third stage were run. Consequently, four regression models were examined. Interaction effects were tested in SPSS using PROCESS (Hayes, 2013). Significant interactions were followed up by examining the conditional effect of empathy on prosocial tendencies at 1 standard deviation (SD) below the mean, at the mean, and 1 SD above the mean of emotion regulation.

3.3. Results

Bivariate correlations between questionnaire measures of empathy, emotion regulation and prosocial behaviour were examined (see **Table 3.1** for a full list of correlations). QCAE-affective empathy and QCAE-cognitive empathy were both positively associated with prosocial tendencies (r = .36, p< .001 & r = .43, p< .001 respectively) and these correlations were not significantly different (z = -.80, p > .05). ERQ-reappraisal was not significantly correlated with prosocial tendencies (r = .13, p = .24). ERQ-suppression was significantly negatively correlated with prosocial tendencies (r = -.27, p = .006). These two correlations were significantly different (Z = 2.69, p < .05).

Table 3.1. Correlations between questionnaire measures

	QCAE-CE	QCAE-AE	PTM total	ERQ-reappraisal
QCAE-AE	.417**			
PTM total	.433**	.358**		
ERQ-reappraisal	.333**	.173	.113	
ERQ-suppression	-.360**	-.529**	-.266**	-.089

Abbreviations: QCAE-AE, Questionnaire of Cognitive and Affective Empathy Affective Empathy subscale; QCAE-CE, Questionnaire of Cognitive and Affective Empathy Cognitive Empathy subscale; ERQ, Emotion Regulation Questionnaire; PTM, Prosocial Tendencies Measure.

* p<.05

** p<.01

To examine whether the associations between the affective and cognitive empathy components and prosocial behaviour were explained by joint variance between the two components or whether they uniquely predicted prosocial tendencies I ran an additional multiple regression analysis. There were unique associations between each empathy component and prosocial tendencies (affective empathy, t = 2.29, p = .024; cognitive empathy, t = 3.67, p < .001) suggesting that these components represented both overlapping and unique predictors of prosocial tendencies.

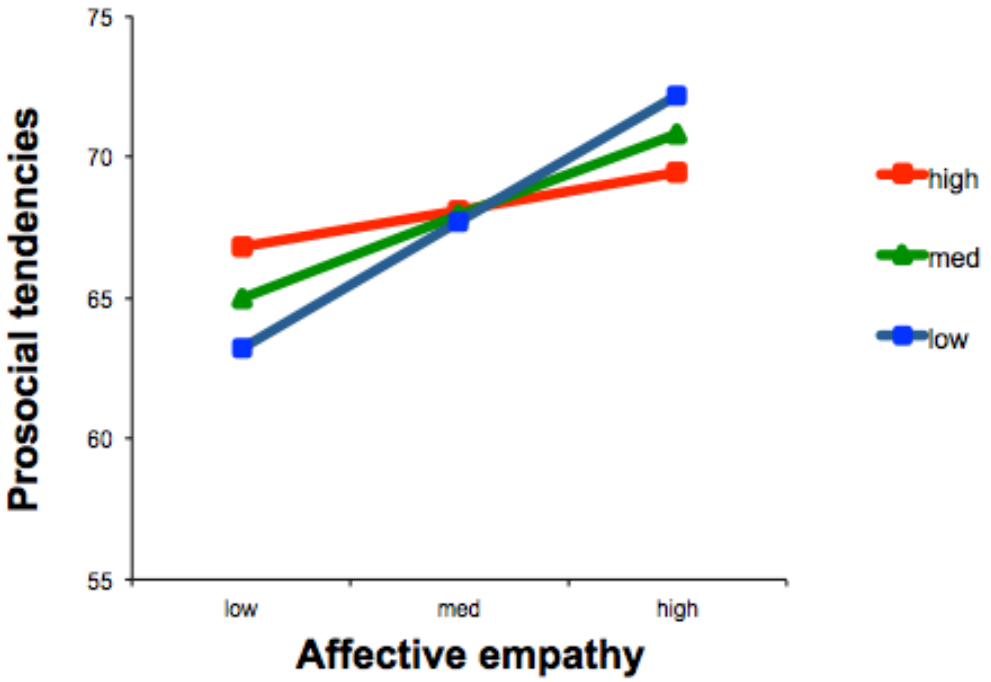
For the first regression model I entered QCAE-affective empathy (first stage), ERQ-reappraisal (second stage), and their interaction term [QCAE-affective empathy x ERQ-reappraisal] (third stage) as predictors of prosocial tendencies. This analysis revealed a significant positive association between QCAE-affective empathy and prosocial tendencies (t = 3.98, p < .001) but not between reappraisal and prosocial tendencies (t = .57, p = .570). Interestingly, the interaction between QCAE-affective empathy and ERQ-reappraisal was significant (t = -2.39, p = .019). At 1 SD below the mean on ERQ-reappraisal there was a significant positive association between QCAE-affective empathy and prosocial tendencies (t = 4.56, p < .001). There was also a significant association at the mean (t = 3.27, p = .002). However at 1 SD above the mean on ERQ-reappraisal the association between QCAE-affective empathy and prosocial tendencies was non-significant (t = 1.08, p = .282) (see **Figure 3.1**).

The significant moderation effect of emotion regulation on affective empathy was also seen when each of the subscales of the affective empathy component (emotion contagion, proximal responsivity, peripheral responsivity) were examined in separate models (all ps < .05).

In other words, affective empathy was associated with prosocial behaviour for those with low and average levels of cognitive reappraisal (with the steepest slope for individuals with lowest level of cognitive appraisal), but those with high levels of cognitive reappraisal presented similar levels of prosocial behaviour regardless of level of affective empathy.

For the second regression model, QCAE-cognitive empathy, ERQ-reappraisal and their interaction term were entered as predictors of prosocial tendencies. This analysis showed a significant positive association between QCAE-cognitive empathy and prosocial tendencies (t = 5.00, p < .001) but not between reappraisal and prosocial tendencies (t = -.39, p =. 699). The interaction between QCAE-cognitive empathy and ERQ-reappraisal was not significant (t = -1.18, p = .243). This pattern of findings suggests that QCAE-cognitive empathy was positively associated with prosocial tendencies regardless of level of reappraisal emotional regulation strategies.

Figure 3.1. moderation of empathy and prosocial tendencies by different levels of emotion regulation.



I also examined the interaction between the two QCAE subscales and ERQ-suppression and their association with prosocial tendencies. These two regression models showed again that both QCAE-AE and QCAE-CE were positively associated with prosocial tendencies (t = 3.98, p < .001 and t = 5.00, p < .001) but that ERQ-suppression was not significantly associated with prosocial tendencies in either model (t = -1.00, p = .32 and t = -1.36, p = .18). Neither of the interactions between QCAE-affective empathy or QCAE-cognitive empathy and ERQ-suppression was significant (both ps >.05).

3.4. Discussion

This chapter investigated associations between empathy and prosocial behaviour, and whether different types of emotion regulation strategy moderate associations between empathy and prosocial behaviour. I found that both

affective and cognitive components of empathy were positively and uniquely associated with self-reported prosocial behaviour. Cognitive reappraisal, but not expressive suppression, played a role in moderating the association between empathy and prosocial behaviour. Specifically, level of cognitive reappraisal moderated the relationship between affective empathy and prosocial behaviour.

The finding that both affective and cognitive empathy are associated with prosocial behaviour supports previous studies suggesting that empathy is a key motivating factor for prosocial behaviour (e.g. Batson, 1998; Carlo et al., 2003; Eisenberg et al., 2010; Eisenberg et al., 1989; Hoffman, 2001). Interestingly, associations between affective and cognitive empathy and prosocial behaviour were not significantly different. Additional analyses showed that cognitive and affective empathy uniquely predicted prosocial behaviour, suggesting that both empathy components play a role in motivating prosocial behaviour. Consequently, whilst it is likely that these two components will often work together in everyday life as they are moderately correlated (e.g. **Chapter 2**; Reniers et al., 2011), this finding raises the possibility that having high levels of just one component could motivate prosocial behaviour, but this needs to be investigated further.

I also observed that expressive suppression was negatively associated with prosocial tendencies. This pattern fits with previous studies suggesting that expressive suppression is a maladaptive emotion regulation strategy (Gross & John, 2003; Gross & Levenson, 1993; John & Goss, 2004; Soto et al., 2011). My results extend these findings by suggesting that in, addition to being related to greater experience of negative emotions, avoidance of close relationships and reports of less life satisfaction (Gross & John, 2003; Gross & Levenson, 1993; John & Goss, 2004; Soto et al., 2011), expressive suppression is also associated with less self-reported prosocial tendencies.

The type of emotion regulation strategy individuals reported to use was important for moderating associations between empathy and prosocial tendencies. Whilst cognitive reappraisal moderated associations between affective empathy and prosocial behaviour, expression suppression did not. In addition, the degree of emotion regulation interacted with the degree of

empathy to predict prosocial behaviour. Affective empathy was positively associated with prosocial behaviour for participants at low and average levels of cognitive reappraisal. This positive association was not evident in participants who reported a high tendency to reappraise. Instead, these individuals had similar levels of prosocial tendencies regardless of level of affective empathy.

Consequently, although empathy is generally assumed to have a significant positive association with prosocial behaviour (Eisenberg et al., 2010; Hoffman., 2008), this may not be the case for all aspects of empathic processing. This finding suggests that affective empathy is an important motivating factor for prosocial behaviour only for particular individuals, which fits with accounts considering a multitude of factors involved in motivating prosocial behaviour (Penner et al., 2005). One explanation is that those with high tendency to reappraise are (at least according to their self-report) more able to change their strategy and viewpoint when evaluating the situation at hand. This capacity may allow one to more readily deduce the desirability of prosocial behaviour even without the experience of the affective components empathy. Whilst I observed a significant moderation of cognitive reappraisal on the association between affective empathy and prosocial behaviour, moderation effects were not evident for associations between cognitive empathy and prosocial behaviour. This lack of association could be because of the overlap in processes involved in cognitive empathy and those involved in cognitive reappraisal. Indeed self-reports of cognitive empathy and cognitive reappraisal were positively correlated in this sample. Processes such as shifting perspective or attention are common to both cognitive empathy and reappraisal. In terms of increasing prosocial behaviour in those individuals high in reappraisal, it is possible that promoting cognitive empathy might elevate the motivation of these individuals to behave prosocially.

Interestingly, I also found that those with the highest levels of self-reported prosocial behaviour were individuals low in reappraisal but high in affective empathy. Given that cognitive reappraisal is positively related to interpersonal functioning (Gross & John, 2003; Gross & Levenson, 1993; John & Gross, 2004; Soto et al., 2011) and prosocial behaviour is generally seen as a positive aspect of interpersonal functioning this result may seem somewhat surprising.

In addition, the model proposed by Eisenberg & Fabes (1992) suggests that those high in experiences of emotional intensity and low in emotion regulation would not manage appropriate prosocial responding and might even display antisocial/aggressive behaviour in response to emotional arousal. However, it has been suggested that high levels of prosocial and altruistic behaviour are not always beneficial and there are cases when acts that are subjectively prosocial can be, to the observer, objectively unhelpful (Oakley, 2013). Future research needs to determine whether the self-reported prosocial behaviours by individuals with high affective empathy and low cognitive appraisal capacities are perceived as objectively helpful/prosocial by the observer. Items on the prosocial tendencies questionnaire assess the self-reported tendency to engage in prosocial behaviours, rather than the quality of them. Future studies could include experimental and/or observational measures to examine this. The types of prosocial responses of individuals high in affective empathy and low in cognitive reappraisal could be compared to those high in cognitive reappraisal and high in affective empathy.

Another promising avenue for future research is to investigate empathy components and emotion regulation strategies in tandem in clinical populations that are thought to show atypical empathy and emotion regulation. For example, autism spectrum disorders, psychopathy and alexithymia have all been associated with both atypical empathy and emotion regulation (Schipper & Petermann, 2013; Swart, Kortekaas, & Aleman, 2009). Finally, the role of empathic concern, i.e. sympathy, in motivating prosocial behaviour has recently been studied theoretically by mathematical models (Szolnoki, Xie, Wang, & Perc, 2011; Szolnoki, Xie, Ye, & Perc, 2013). These models suggest that the development of empathic concern can lead to development of cooperation in economic games (termed evolutionary games by the authors). Consequently, such models suggest potential mathematical principles that could be applied in future studies to model how empathy might lead to prosocial behaviour, which is explored in further detail in **Chapter 6**.

3.5. Conclusions

Overall, these findings suggest that both affective and cognitive empathy are motivating factors for prosocial behaviour. However, empathy and emotion regulation can also interact to predict different levels of self-reported prosocial behaviour such that there is not always a significant positive association between affective empathy and prosocial behaviour. These results could help to account for why associations between empathy and prosocial behaviour can sometimes be modest or weak. These results also suggest that further investigations of the type of prosocial behaviours exhibited by individuals with varying levels of empathy and emotion regulation could be relevant as we try to understand how empathy might motivate prosocial ways of interacting with others.

In the next chapter, **Chapter 4**, I describe an fMRI study that further probes how empathic processes relate to social behaviour by examining how children with low levels of empathy and high levels of antisocial behaviour process information about other people's pain.

CHAPTER 4: Chapter background

In the previous chapter I examined how empathy may motivate positive social behaviours, that is, prosocial behaviour. In the current study, I extend this aim to focus on how individual differences in empathy relate to negative social behaviour, that is, antisocial behaviour.

Children with conduct problems display high levels of antisocial behaviour and incur a considerable economic and social cost. These children also display atypical empathic responses to others' distress, which may partly account for their violent and antisocial behaviour. In children with conduct problems, callous traits index a lack of empathy and confer risk for psychopathy in adulthood. Investigating neural responses to images of other people in pain can be used as a proxy measure of empathic processing, yet studies in children with conduct problems have been inconclusive. In this chapter I report a study that used functional magnetic resonance imaging (fMRI) to examine neural responses to images of other people in pain in a large sample of children with conduct problems and varying levels of callous and unemotional traits, and a control group matched for IQ, age, socioeconomic status and ethnicity. I also acquired parent and teacher ratings of conduct problem symptoms and callous and unemotional traits to examine individual differences in neural response. I hypothesised that (1) children with conduct problems would show reduced neural responses to other peoples pain, compared to controls, in key regions of the brain previously associated with affective and social information processing (anterior insula, anterior cingulate cortex, inferior frontal gyrus). (2) Variation in activation in these regions would be predicted by conduct problem symptoms and callous traits.

4.1. Introduction

Conduct problems (CP) in children include physical aggression, lying, theft, and cruelty to others. Given that empathy plays a key role in inhibiting aggression and promoting prosocial behaviour (see **Chapter 1**) it has been hypothesised that antisocial behaviours in children with conduct problems may reflect atypical empathic responses to other people's suffering (Blair, 2005). While there is no clinical diagnosis of psychopathy in childhood, there is abundant evidence that psychopathic traits and behaviours can be observed in children. In childhood, particular subsets of children with conduct disorder have elevated levels of psychopathic traits, which are termed callous-unemotional traits in research studies and 'limited prosocial emotions' in the new DSM-5 guidelines. Callous-unemotional traits in children can persist into adulthood (Lynam et al., 2007) and are highly heritable (Viding et al., 2005). In contrast, antisocial behaviour in children without callous-unemotional traits appears to be primarily driven by environmental influences (Viding et al., 2005). Heterogeneity in empathy at the behavioural level in children with CP is well captured by a questionnaire assessment of callous-unemotional (CU) traits (Kimonis et al., 2008).

One method for investigating neural processing that may be associated with empathy is to measure the perception of other people in pain. Delineating these responses in children with CP is of particular interest given that this group often inflicts pain on others (Romeo, Knapp, & Scott, 2006). fMRI studies in healthy populations have identified a network of brain regions activated by the observation of pain in others. This network includes regions linked to sensory processing such as somatosensory cortices, regions linked to processing social and affective information, such as anterior insula (AI) and anterior cingulate cortex (ACC) (Fan et al., 2011; Lamm et al., 2011; Singer & Lamm., 2009) and regions linked to cognitive reappraisal such as the inferior frontal gyrus (IFG).

Atypical function and structure in several of these regions, including AI, ACC and IFG have been implicated in the pathophysiology of childhood CP and adult psychopathy (Anderson & Kiehl, 2012). For example, reduced grey matter

volume and atypical function of AI has been reported in children with conduct problems and adults with psychopathy (Anderson & Kiehl, 2012; Sebastian, et al., 2012; Sterzer, Stadler, Poustka, & Kleinschmidt, 2007). The ACC is of particular interest in the current thesis since, as highlighted in **Chapter 1**, this region is well known for its role in social behaviour (Apps, Lockwood, et al., 2013; Behrens et al., 2008; Rushworth, Behrens, Rudebeck, & Walton, 2007) and lesions to this region appear to disrupt the value assigned to social stimuli (Rudebeck et al., 2006). Moreover, individuals with psychopathy have been shown to display a reduced error related negativity, putatively sourced in the ACC, when observing others outcomes during a social interaction (Brazil et al., 2011). Recent studies also indicate that grey matter volume and activity in the ACC correlates with psychopathic and callous traits (Anderson & Kiehl, 2012; Cope et al., 2012; De Brito et al., 2009).

However, to date, only two studies have investigated processing of other people's pain in children with CP, with inconclusive results. Decety and colleagues found that, compared with controls, children with CP showed increased neural responses to others in pain in regions including the insula, anterior midcingulate, striatum and amygdala (Decety, Michalska, Akitsuki, & Lahey, 2009). Aggressive CP symptoms were positively correlated with inferior frontal gyrus, cingulate cortex, amygdala and periaqueductal grey responses. Although these findings are of interest, CU traits were not measured and the CP sample was small (N=8), making replication and extension of this work important. Another study measured event-related potentials, and found frontal N120, thought to reflect early affective arousal; and central-parietal late-positive potentials (LPPs), thought to index reappraisal of unpleasant stimuli (Cheng, Hung, & Decety, 2012; Fan & Han, 2008) were reduced in children with CP and high levels of CU traits relative to TD children when viewing pictures of others in pain (Cheng et al., 2012). Findings from these two studies provide preliminary evidence that children with CP show atypical responses to others' pain, which may be partially driven by CU traits.

4.1.1. *The current study*

In the current chapter I compared neural responses to others' pain in children with CP and TD controls using fMRI. I also explored heterogeneity of neural responses within the CP group by studying unique dimensional contributions of CU traits and CP symptoms. The standard research measure of CU includes scales that characterise callous, uncaring, and unemotional attributes (Essau et al., 2006). I was particularly interested in the effect of callous traits, since these index reduced empathy, guilt and remorse for others' suffering; whilst uncaring traits relate to a lack of concern about task performance and unemotional traits relate to an absence of emotional expression (Essau et al., 2006).

4.1.2. *Hypotheses*

On the basis of previous research suggesting reduced empathy in children with CP (Blair, 2005; Jones et al., 2010; Schwenck et al., 2012), I predicted reduced neural responses to Pain (relative to No Pain) in the CP compared to TD group in regions previously shown to respond in empathy for pain paradigms that all thought to be atypical in CP, including the AI and ACC and IFG. Within the CP group, I also predicted that CU traits (in particular the 'callous' subscale) might be negatively related to neural responses to others' pain, while CP symptoms might show a positive relationship (Sebastian, McCrory, et al., 2012). I focused on unique variance associated with CU traits and CP symptoms (i.e. controlling for one another), as several lines of evidence suggest that these variables exert suppressor effects on one another (Hicks & Patrick, 2006; Sebastian et al., 2012).

4.2. Materials and Method

4.2.1. *Participants*

Right-handed boys aged 10-16 (mean (SD): controls=13.68 (1.68); CPs=14.05 (1.69)) were recruited from the community via advertisements and local schools. One-hundred-and-forty-three parents and teachers completed screening questionnaires. CP was assessed using the Child and Adolescent Symptom Inventory (CASI-4R) (Gadow & Sprafkin, 2009) Conduct Disorder scale (CASI-CD). CASI-CD cut-off scores for inclusion in the CP group were: Parent report=4+ (ages 10-12) 3+ (ages 12-16) or Teacher report=3+ (ages 10-12) 4+ (ages 12-14) 6+ (ages 15-16). These scores are associated with a clinical diagnosis of CD (Gadow and Sprafkin., 2009). CU traits were assessed using the Inventory of Callous-Unemotional Traits (ICU) (Essau et al., 2006). Total scores for the three ICU subscales (callous, uncaring and unemotional) were calculated. Both CASI-CD and ICU were scored by taking the highest ratings from either the parent or teacher questionnaire for each item (Piacentini, Cohen, & Cohen, 1992). For two children with CP only parent ratings were available. The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) was used to screen for psychopathology (hyperactivity, CP, emotional symptoms, peer problems) in the controls. All control participants scored below the CP group median on the ICU and in the normal range on the CASI-CD and SDQ. For both groups, exclusion criteria included previous diagnosis of neurological or psychotic disorder, including autism spectrum disorders, and current prescription for psychiatric medication (all children were unmedicated). Written informed consent from parents and written assent from participants was obtained.

Fifty-eight children were scanned (39 CPs, 19 controls), with usable data from 37 CPs and 18 controls. Exclusions were due to: scanner refusal (1 CP), and teacher questionnaire data obtained after scanning indicating that the child no

longer met group criteria (1 CP, 1 control). Groups were matched on IQ, age, ethnicity and socioeconomic status (**Table 4.1**).

Table 4.1. Participant characteristics

	Group		
Characteristics and questionnaires	Controls (N=18)	CP (N=37)	P Value[a]
Demographic variables			
Age[b]	13.68 (1.68)	14.05 (1.69)	.456
Socioeconomic Status[b]	3.07 (1.01)	3.23 (1.26)	.635
F-IQ[c,d]	102.83 (11.69)	101.17 (13.46)	.656
V-IQ[c,d]	53.06 (8.73)	49.92 (10.96)	.295
P-IQ[c,d]	49.67 (8.61)	50.33 (9.57)	.804
Ethnicity[b,e]	13: 2: 2: 1	26: 3: 6: 2	1.00
Inventory of Callous-Unemotional Traits[f]	24.17 (4.85)	42.97 (10.67)	<.001
Child and Adolescent Symptom Inventory			
Conduct Disorder[f]	.56 (.70)	10.14 (6.18)	<.001
ADHD[d,g]	9.47 (7.47)	25.04 (13.75)	<.001
Generalised Anxiety Disorder[d,g]	2.71 (3.07)	7.46 (5.37)	.001
Major Depressive Episode[d,g]	2.61 (1.09)	6.38 (5.40)	.005
Alcohol Use and Disorders[c,d]	1.22 (1.99)	2.11 (3.88)	.366
Drug Use and Disorders[c,d*]	.17 (.51)	2.50 (6.62)	.143

Abbreviations: CP, Conduct Problems; F-IQ, Full IQ score from the 2-subtest Weschler Abbreviated Scale of Intelligence; V-IQ, Verbal IQ score; P-IQ, Matrix reasoning IQ score; ADHD, attention-deficit/hyperactivity disorder.

[a]All *p*-values obtained using *t*-tests except for Ethnicity (Fisher's exact tests used)

[b]Measures taken at screening phase - parent report

[c]Child at scanning session

[d]Missing data from 1 CP

[e]Ethnicity: White:Black:Mixed:Asian

[f]Meaures taken at screening phase - parent and teacher report

[g]Measures taken at scanning session - parent report

*The Drug Use and Disorders scale requires participants to rate the frequency of any substance use on a 5-point scale from never' to 'almost daily'. The list of drugs includes cannabis, amphetamines, cocaine, opiates, hallucinogens, solvents, GHB as well as medicines used in an abusive way.

4.2.2. *Procedure*

Stimuli were 192 digital photographs showing another person's hand or foot in painful or non-painful situations (Gu et al., 2010) 'Pain' and 'No Pain' stimuli (96 pictures per condition) were matched on physical properties, and were validated as eliciting activations in the hypothesised regions in a previous study (Gu et al., 2010). Stimuli were presented in pain and no pain blocks lasting 20 seconds and consisting of 8 images, each displayed for 2000ms with a 500ms ISI. Blocks were pseudorandomised, with the same block type never presented more than twice in a row. A fixation cross was presented for 15 seconds every 6 blocks.

To ensure attention, participants performed a hand/foot key press judgment on every trial. Participants practiced outside the scanner with painful and non-painful images not seen in the main experiment, until ≥80% accuracy was reached.

4.2.3. *Psychometric and questionnaire measures*

Participants completed the Weschler Abbreviated Scale of Intelligence two-subtest version (Wechsler, 1999), and the Alcohol/Drug Use Disorders Identification Tests (Berman, Bergman, Palmstierna, & Schlyter, 2005). A parent or guardian completed the CASI-4R scales for symptoms commonly co-morbid with CP, including ADHD, Generalized Anxiety Disorder, and Major Depressive Episode (**Table 4.1**).

4.2.4. *fMRI data acquisition and analysis*

A Siemens Avanto 1.5-T MRI scanner was used to acquire 189 multislice T2*-weighted echo planar volumes with blood oxygenation level–dependent contrast (1 run of 9 minutes). The sequence was based on (Weiskopf, Hutton, Josephs, & Deichmann, 2006). Functional sequence acquisition parameters were as follows: 35 2mm slices acquired in an ascending trajectory with a 1mm gap; TE=50ms; TR=2975ms; slice tilt=−30° (T>C); flip angle=90°; field of view=192 mm; matrix size=64x64. A 5.5-minute T1-weighted MPRAGE scan was acquired for coregistration, normalization and overlay, and fieldmaps were acquired for unwarping. Data were analyzed using Statistical Parametric Mapping (SPM8; Weiskopf et al., 2006). The preprocessing pipeline was as follows: the first 5 and last 2 volumes were discarded. Data were then realigned, unwarped using a fieldmap, normalized via segmentation of participants' structural scans, written with a voxel size of 2x2x2mm, and smoothed with a 8mm Gaussian filter.

Given the potential for motion artifacts using a developmental sample, I followed a number of procedures to limit any influence of motion on the collected data. Firstly, participants were shown a short slide show of brain scans taken from previous participants who had either stayed still or moved varying amounts. Participants were also given a short practice/localizer scan, after which feedback was given on how still they kept. After estimation of the realignment parameters, I ran a script to search for motion of more than 1mm (x,y and z directions) or 1 degree (pitch, roll, yaw) in any direction between acquisition of one volume and the next. Volumes flagged by the script (as well as surrounding volumes) were then inspected visually for motion artifacts. For a few random participants, the whole time series was inspected for motion artifacts to check the validity of the threshold chosen in the script. On the basis of previous studies from our laboratory I decided a priori to exclude any participants where more than 10% of the volumes were corrupted by motion artifacts. No participants reached this threshold.

For 14 participants (11 CPs, 3 controls), extra regressors were included within the first-level analysis design matrix to model any images corrupted by motion. These images were removed and the adjacent images interpolated to prevent distortion of the between-subjects mask. Data were high-pass filtered at 128 seconds to remove low-frequency drifts. For one CP, half of the fMRI time series (91 scans) was discarded due to the participant falling asleep midway through scanning.

4.2.4.1. First-level and second-level analysis

After standard preprocessing, a block analysis compared neural activity associated with pain and no pain conditions. Regressors included Pain and No Pain (blocks of 20 seconds duration) and fixation (15 seconds), modeled as boxcar functions convolved with a canonical hemodynamic response function. The 6 realignment parameters were also modeled as effects of no interest. At the first level, Pain>No Pain and No Pain>Pain contrasts were created. Contrast images were entered into second-level analyses, where group (CP vs. control) served as a between-subjects variable in independent-samples *t* tests. Main effects are reported at $p < .05$, family-wise error (FWE) corrected across the whole brain (See **Appendix 2**); while regions from a whole brain analysis showing a condition x group interaction are presented at $p < .005$, $k \geq 10$, uncorrected in (See **Appendix 2**) (no interaction results survived FWE-correction across the whole brain). I investigated the condition by group interaction in three a priori regions of interest (bilateral AI, ACC and IFG). ROIs were anatomically defined using masks from the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). The MarsBaR toolbox (http://marsbar.sourceforge.net) was used to calculate mean contrast estimates across bilateral ROIs. Group differences were assessed at a standard statistical threshold of $p < .05$ (Eisenberger et al., 2010; Masten et al., 2011).

I selected the anterior portion of the insula by modifying the aal atlas mask to include all voxels y>0 on the basis of several previous studies suggesting that it

is the anterior portion of the insula that is involved in affective-motivational aspects of vicarious pain processing (Fan et al., 2011; Lamm et al., 2011). For ACC I used the standard ACC aal mask. The peak co-ordinate from Lamm et al. (Lamm et al., 2011) falls within this mask in the left hemisphere and falls on the border of the mask in the right hemisphere. I included hemisphere as a separate factor in repeated measures ANOVAs for each ROI. These showed no significant differences between hemispheres for the condition*group interaction effect of interest for any ROI (ps>.40). Since I had no a priori hypothesis regarding laterality, I therefore collapsed data across bilateral ROIs.

I also performed additional analysis given the particular interest of the role of the ACC in social cognition in this thesis. The aal ACC mask includes a number of sub-regions including BA24, BA32 and BA25 (Vogt et al., 1995). Given that previous studies have suggested the ACCg is particularly sensitive to social information (as compared to other portions of the ACC), I took the peak co-ordinate from 4 studies showing activation in the ACCg to social information processing (Apps, Green, et al., 2013; Apps & Ramnani, 2014; Behrens et al., 2008; **Chapter 5**) and averaged these coordinates to conduct a small volume correction using a 6mm sphere after initial thresholding of p<.005 uncorrected.

4.3. Results

4.3.1. *Behavioural data*

Mean reaction times (RTs) and percentage error rates were calculated. For mean RTs, a group (CP vs. control) by condition (pain, no pain) ANOVA showed no main effect of group ($F(1,53) = .02$, $p = 0.89$) but a main effect of condition ($F(1,53) = 71.85$, $p < .001$) with significantly slower RTs when classifying hands and feet in the pain condition (910.08, SD=140.15) compared to no pain (862.72, SD = 129.73). There was no interaction between group and condition

(F(1,53) <.001, p > 0.99). Error data showed a marginal main effect of group (F(1,53) = 3.17, p = 0.08) with a trend for more errors in the CP than control group, and a main effect of condition (F(1,53) = 6.40, p = 0.014) with significantly more errors when classifying hands and feet in pain compared to no pain (6.82, SD = 5.05 vs. 5.63, SD = 4.55). There was no group by condition interaction (F(1,53) = .061, p = 0.81).

4.3.2. *fMRI data: main effect and group x condition interaction*

Results from whole brain analyses for the main effect of Pain>No Pain (and the reverse), and the group by condition interaction are displayed in **Appendix 2**. Main effects were found in regions previously associated with empathy for pain, and largely replicated a previous study using the same stimuli (Gu et al., 2010). ROI analyses for Pain>No Pain revealed the predicted pattern of reduced response in the CP relative to control group in bilateral AI (t(53)=2.08, p = 0.02), ACC (t(53)=1.66, p = 0.05) and IFG (t(53)= 2.45, p <0.01). Results from the additional analysis of the average ACCg coordinate also showed a significant group x condition interaction effect (MNI coordinates x=0 y=20 z=24, k=14, z=2.8, p < .05, FWE-SVC).

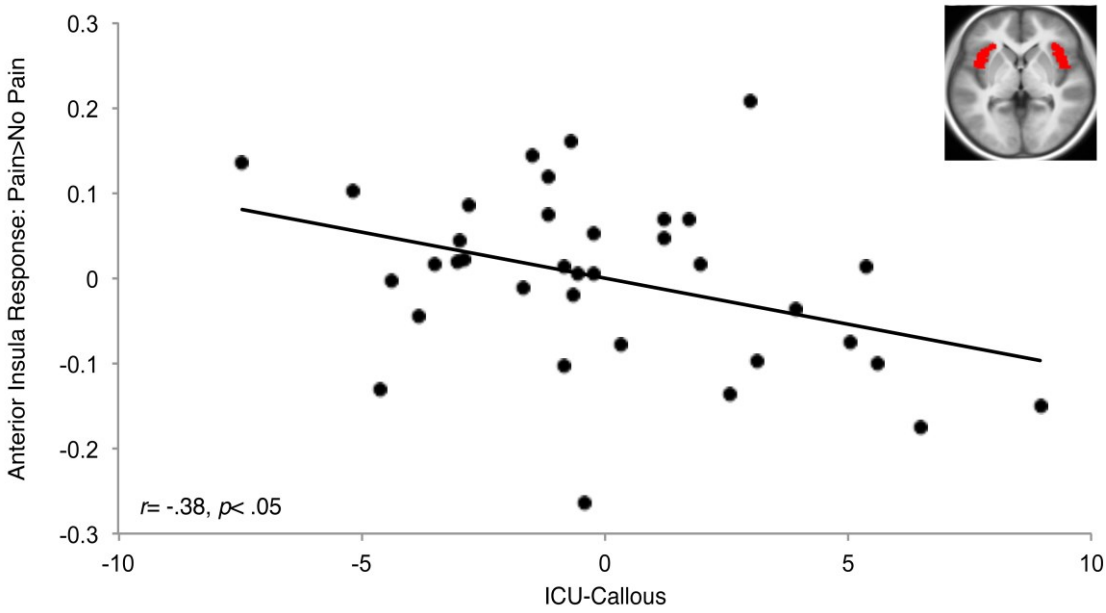
4.3.3. *fMRI data: associations with ICU and CP symptoms*

I then examined the second hypothesis, that callous traits would be associated with reduced ROI responses to Pain>No Pain within the CP group. On the basis of previous findings showing that CP symptoms and CU traits exert suppressor effects on one another (see Hicks & Patrick, 2006; Sebastian, et al., 2012) I conducted multiple regressions to investigate unique contributions of ICU subscales (callous, uncaring, unemotional) and CP symptoms to neural responses in the ROIs. One participant was excluded from these analyses due

to missing data.

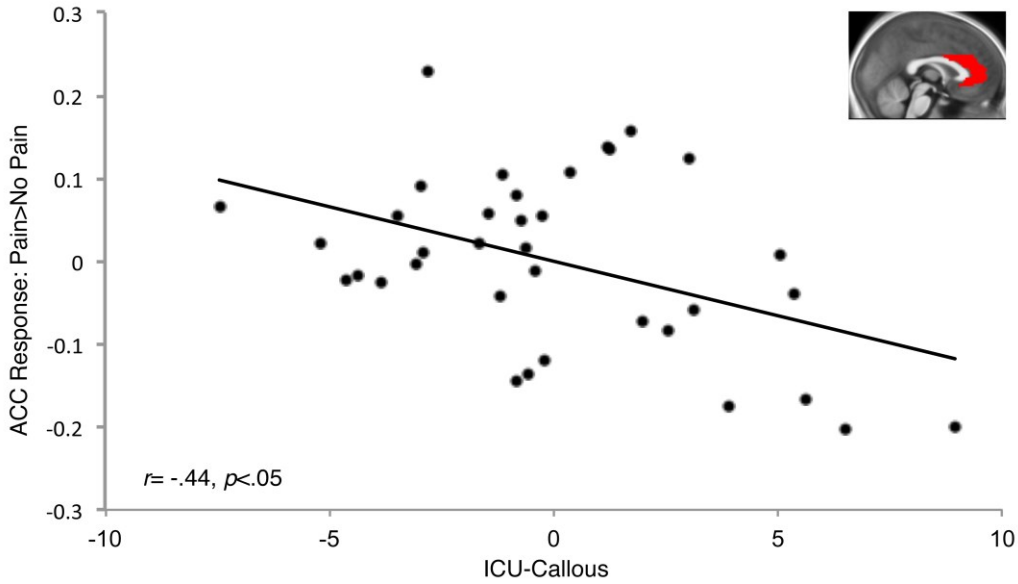
In AI, a significant negative relationship was found between unique variance associated with callous traits and neural response (β = -.625, p = 0.03) (**Figure 4.1**). Neither CP symptoms nor uncaring or unemotional subscales were associated with AI response (all *ps* >.10). In ACC, a significant negative relationship was found between unique variance associated with callous traits and neural response (β = -.729, p = 0.01) (**Figure 4.2**), while a positive relationship was found between unique variance associated with CP symptoms scores and neural response (β = .485, p = 0.02) (**Figure 4.3**). No relationships were found in relation to the uncaring or unemotional subscale scores (*ps* > 0.24). In IFG, no associations with unique variance were found. To investigate potential effects of commonly co-morbid attention-deficit hyperactivity, generalized anxiety and depression symptoms, I included these variables in follow-up regression analyses. All significant results remained at $p < 0.05$ and none of these variables predicted AI or ACC response (all *ps* > .25). In addition, when age was included in follow-up regression analyses all results remained significant at $p < 0.05$.

Figure 4.1. Partial regression plot for the CP group (n=36), showing a negative association between bilateral AI response to Pain>No Pain, and unique variance associated with ICU-callous traits after controlling for CASI-CD, ICU-unemotional and ICU-uncaring scores.



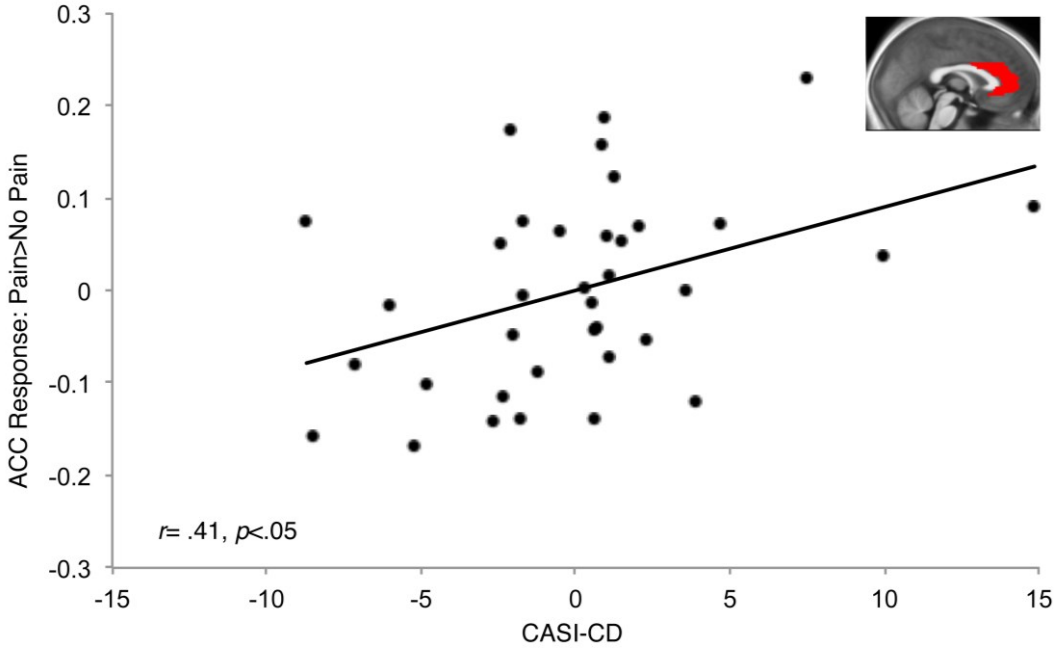
Notes: (*Inset*) Horizontal section (z=0) of bilateral AI ROI overlaid on an average T1 structural image from all participants. Bilateral AI response was calculated by averaging left and right AI response. *P* and r reflect partial correlation coefficients.

Figure 4.2. Partial regression plots showing associations with ACC response to Pain>No Pain in the CP group (n=36). Negative relationship between ACC response and unique variance associated with ICU-callous traits, after controlling for CASI-CD, ICU-unemotional and ICU-uncaring scores.



Notes: (*Inset*) Sagittal section (x=0) of ACC ROI overlaid on an average T1 structural image from all participants. Bilateral ACC response was calculated by averaging left and right ACC response. *P* and r reflect partial correlation coefficients.

Figure 4.3. Partial regression plots showing associations with ACC response to Pain>No Pain in the CP group (n=36). Positive relationship between ACC response and unique variance associated with CASI-CD scores after controlling for ICU-callous, unemotional and uncaring subscale scores



Notes: (*Inset*) Sagittal section (x=0) of ACC ROI overlaid on an average T1 structural image from all participants. Bilateral ACC response was calculated by averaging left and right ACC response. *P* and r reflect partial correlation coefficients.

Bivariate correlations were also conducted within the CP group between neural responses to Pain>No Pain in regions of interest (AI, ACC, IFG) and CP symptoms and ICU subscales (callous, unemotional and uncaring traits) to determine if the unique associations between these measures and neural response were not evident in bivariate associations. In the AI and ACC, no significant bivariate relationships were found (all *ps* > 0.07). In IFG, positive correlations were found between neural response and both CP symptoms (*r* = .40, *p* = 0.02), and unemotional traits *(r* = .44, *p* < 0.01). There was no significant R-squared change when adding CP symptoms after unemotional traits, or vice-versa, into a regression model indicating that common variance

between CASI-CD and unemotional traits drove the positive relationship with IFG.

I also extracted the significant peak in the ACCg identified in the additional analysis and explored unique and bivariate associations with CP symptoms and ICU subscales in the CP group. This analysis largely replicated the results from associations with the average response over the entire ACC mask, with unique variance from callous traits being negatively associated with ACCg response (β = -.496, p = .08 [marginal]) and CP symptoms being positively associated (β = .534, p = .01). Uncaring and unemotional traits showed no significant unique associations (all ps > .80) and these results did not change when commonly co-morbid symptomatology and age were added to the analyses (all ps < .10).

4.4. Discussion

In this study, I show reduced neural responses to others' pain in children with conduct problems compared with controls, in three key regions, AI, ACC and IFG, associated with the perception of others pain in previous studies. This is the first fMRI study to investigate the processing of others' pain (as an index of empathic processing) in a large sample of children with CP, using a well-controlled task matched for visual and social content. I also show a negative association between callous traits and AI/ACC responses. Additional analyses to examine whether group differences were apparent in the ACCg also showed reduced neural response to others' pain in the CP group compared to the control group, and diverging associations between callous traits and CP symptoms.

Meta-analyses have indicated that the AI and ACC are consistently activated during the vicarious pain and have a close functional relationship (Fan et al., 2011; Lamm et al., 2011; Medford & Critchley, 2010; Mufson & Mesulam, 1982; Pandya, Van Hoesen, & Mesulam, 1981). The AI is proposed to play an important role in sensory integration and interoceptive awareness (Craig, 2009),

and may be involved in awareness of unpleasant feelings during empathy for pain (Medford & Critchley, 2010). Interestingly, abnormal AI function and structure has frequently been reported in both children with CP and in adults with psychopathy (Decety et al., 2009; Ly et al., 2012; Sebastian et al., 2012; Sterzer et al., 2007). However, the observation of reduced AI response is at odds with one study, which found increased AI response in children with CP (Decety et al., 2009). This could be because in that study pain caused by accident was contrasted with pain caused by others, whereas the pain and no pain conditions were matched for agency. Increased AI reactivity may reflect differences in agency processing rather than pain processing per se. Differences in the samples between the two studies (e.g. levels of callous traits) may also have contributed to the divergent findings. These data provide additional support for the view that atypical AI function represents a neural marker of disrupted empathic processing in CP and that AI hypoactivity relates to differences in processing others' pain.

The ACC is well known for its role in social behaviour (Apps, Lockwood, et al., 2013). Like AI, atypical ACC function in CP has been reported previously, again with mixed findings (Sterzer, Stadler, Krebs, Kleinschmidt, & Poustka, 2005). One study reported reduced ACC response to negative pictures in CP (Sterzer et al., 2005), while another found greater ACC response in children with CP to videos of others in pain vs. no pain (Decety et al., 2009). This finding provides converging evidence that ACC function is atypical in CP and in particular that there is hypoactivity of response during empathy for pain.

To address the second aim I explored dimensional unique contributions of CU traits and CP symptoms to ROI responses. As predicted, unique variance associated with callous traits was negatively related to AI and ACC response. This association was also at trend when examining peak response in the ACCg. Since the callous subscale of the ICU contains items reflecting poor empathy in everyday life, these findings provide evidence of convergent validity between questionnaire and neural measures of empathy in CP. Moreover, callousness is an important feature of adult psychopathy (Blair et al., 2005) and childhood CU traits predict adult psychopathy (Lynam et al., 2007). Blunted neural responses to pain in children with higher levels of callous traits may characterise a

developmental vulnerability to serious antisocial behaviour; for a minority, such a pattern may interact with other vulnerability factors to increase risk of adult psychopathy.

Interestingly, CP symptoms were positively correlated with ACC response, and this same association was evident when examining the ACCg. These results complement recent findings (Sebastian et al., 2012) showing opposing unique contributions of CU traits and CP symptoms to neural response in the amygdala. Heterogeneity in CP may help to explain inconsistencies across previous studies reporting both increased and decreased ACC responses in CP (Decety et al., 2009; Sterzer et al., 2005).

More generally, these data highlight that children with CP are a heterogeneous group with varying neurocognitive vulnerabilities; with callous traits of particular importance in predicting empathic dysfunction.

4.4.1. *Limitations*

Limitations of the current study include the use of a research diagnosis of CP, and a focus on males. Replication in a clinically diagnosed sample will be important, as will investigation of potential gender differences. Additionally, the task did not allow exploration of the function of component processes within the empathy for pain response in CP. Future studies should address whether there is a specific aspect of this response which is atypical in CP, e.g. basic arousal, interoceptive processing, or higher-level emotional responses to others' suffering. Finally, replication and extension of the current study is required. The present paradigm used social pictures as stimuli, rather than examining social interactions with another agent, and therefore it is unclear how the reduced responsiveness to others pain observed in the current studies is apparent during social interactions. Moreover, given that the current paradigm used a block design I cannot directly examine the specific component processes associated with the representation of others' pain, such as pain

prediction/anticipation and the outcome of observing others receive pain. Finally, as the current paradigm does not contain a 'self' pain condition I cannot examine whether the difference in neural responses relate specifically to atypical processing for social information.

4.4.2. *Conclusions*

Despite these limitations, these data extend understanding of the neural basis of CP and empathy in several important ways. To my knowledge, this is the first study to investigate empathic pain processing in a large sample of children with CP compared with controls on a task matched for visual and social content. I show reduced neural responses to others' pain in children with CP. Second, I show that callous traits in particular may underlie atypical neural responses to others' pain in CP, which may represent an early neurobiological marker for later psychopathy. Third, the finding that callous traits and CP symptoms show opposing relationships with ACC response suggests a potential explanation for mixed reports of hyper-activation (Decety et al., 2009) and hypo-activation (Sterzer et al., 2005) of ACC to negative affective stimuli in CP. Clinically, these data may have consequences for empathy training implementation (e.g. in relation to victim empathy) in children with high levels of callous traits. Systematic evaluation of training outcomes should take callous traits into account. It remains an empirical question whether empathic responding can be normalised in children with CP (and varying levels of callous traits) or whether behavioural equivalence is better achieved through compensatory strategies which leverage spared cognitive processes (Jones et al., 2010; Schwenck et al., 2012).

In the next chapter, **Chapter 5**, I will present a study investigating neural responses to others' positive outcomes, namely reward, and variation with individual differences in empathy that extends the design in the present chapter and addresses some of the limitations (albeit in a healthy adult population).

CHAPTER 5: Chapter background

In the last chapter I examined neural responses to other people's pain in children with low levels of empathy. In the current study, I extend this aim to focus on how individual differences in empathy underpin neural responses to others positive experiences, in particular other peoples' reward. Successful empathising can depend on the ability to predict when others are likely to receive rewards. Yet, whilst a plethora of research has examined the neural basis of predictions about the likelihood of receiving rewards ourselves, very little is known about the neural mechanisms that underpin variability in *vicarious* reward prediction. Human neuroimaging and non-human primate studies suggest that a sub-region of the anterior cingulate cortex in the gyrus (ACCg) is engaged when others receive rewards. Does the ACCg respond when we observe others about to receive a reward and does this response vary with individual differences in trait empathy? In the present study I used fMRI to examine neural responses to cues that signalled the likelihood of others reward in a sample of adult males who were matched with a social confederate. I also measured individual differences in empathy. I hypothesised that (1) the ACCg is engaged when predictions are made about the probability of another person receiving a reward and (2) the extent to which the ACCg is specialised for processing others' rewards is positively associated with trait empathy as measured by the QCAE.

5.1. Introduction

The successful prediction of future rewards is fundamental for adaptive behavior. The neural mechanisms that underpin reward prediction for oneself are becoming increasingly well understood (Schultz, 2013). However, during social interactions, stimuli are often predictors of rewards for others, not exclusively ourselves. Effectively cooperating, competing or empathising with another requires the ability to compute the value of stimuli that predict rewards for others (Ruff & Fehr, 2014). Yet, very little is known about how vicarious reward predictions are processed in the brain. Moreover, there is a dearth of knowledge regarding how individual differences in social functioning are related to neural response to others' reward.

The dorsal anterior cingulate cortex (dACC) signals predictive information about reward value, including the probability and magnitude of future rewards (Rogers et al., 2004; Sallet et al., 2007; Shidara & Richmond, 2002). This region is also engaged when processing social information (Behrens et al., 2008; Gabay, Radua, Kempton, & Mehta, 2014; Lamm et al., 2011). Recently, a model of the dACC was proposed that unifies these different facets of its function (Apps, Lockwood, et al., 2013). This model posits that a sub-region of the ACC in the gyrus (ACCg) – lying in the anterior portions of the midcingulate cortex (areas 24a'/24b') (Vogt et al., 1995) - is sensitive to processing information about rewards for other people, including probabilistic predictions about rewards that others are likely to receive (Apps, Lockwood, et al., 2013). Several lines of evidence support this model. First, there are neurons in the ACCg that respond when a monkey views cues that indicate that another monkey will receive a reward (Chang et al., 2013), and in the dACC that respond when monkeys predict the decisions of a conspecific in an economic game (Haroush & Williams, 2015). Second, lesions to the ACCg reduce the value assigned to social stimuli, leaving the processing of non-social stimuli intact (Rudebeck et al., 2006). Third, hemodynamic responses in this region vary with the net-value of rewards received by others, the volatility of social information, predictions about the value of others' actions and predictions of social approval from others

(Apps & Ramnani, 2014; Behrens et al., 2008; Boorman et al., 2013; Jones et al., 2011). Taken together, these studies point to a central role for the ACCg in processing information about others' rewards. Yet, a key untested hypothesis from this model is that the ACCg is engaged when predictions are made about the probability of another person receiving a reward. Therefore the first aim of this chapter was to test this hypothesis.

A second hypothesis derived from this model is that individual differences in social functioning, specifically empathy, vary with the extent to which ACCg is specialised for processing others' rewards. Empathy can be broadly defined as the capacity to understand and resonate with the experiences of others (Singer and Lamm, 2009). Empirical and theoretical accounts have suggested that the ACC is involved in empathising (Engen & Singer, 2013; Lamm et al., 2011), but prior work has largely focused on response of this region to processing others' pain and other negative outcomes (reviewed in Lamm et al., 2011), rather than positive, rewarding outcomes. The propensity to feel empathy varies substantially between individuals (Bird & Viding, 2014; Blair, 2005; **Chapter 2**) but the mechanisms that underpin individual differences in vicariously processing another's rewards are still relatively poorly understood. Therefore the second aim of this chapter was to test the hypothesis that the extent to which the ACCg is specialised for processing others' rewards is positively associated with trait empathy.

5.2. Materials and method

5.2.1. *Participants*

Thirty-two right-handed healthy males (age 19-32, M=22.7 SD=3.0) were recruited through university participant databases. Exclusion criteria included previous or current neurological or psychiatric disorder, non-normal or non-

corrected to normal vision, non-native English language and previous or current study of psychology. This latter criterion was employed due to concerns that prior experience of studying psychology could compromise participants' belief in the deception used in the protocol. Two participants were excluded from the analysis (one due to excessive motion (> 10% of scans) and one due to neurological abnormalities) leaving a final sample of 30. All participants gave written informed consent and the study was approved by the local departmental research ethics committee.

5.2.2. *Experimental task*

5.2.2.1. Design

I examined the processing of cues that signalled the probability with which a first-person and a third-person would receive a reward. A 2 x 2 factorial design (agency (self vs. other) and probability (high 80% vs. low 20%)) was employed to examine activation time-locked to the cues (see figure 1).

On each trial during the experiment participants saw cues that indicated the probability with which they (first-person, or 'self') or the other participant (third-person, or 'other') were likely to win points. These cues were represented as pie charts in order to depict the level of probability explicitly and minimise any requirements for reward learning across the task. The cues for 'self' and 'other' differed in colour but were luminance matched. 'Self' cues had the word 'YOU' written above them whilst 'other' cues had the name of the other participant (a confederate) written above them. This ensured that participants were explicitly aware of whether the cues predicted outcomes for themselves or for the other participant.

Following the cue an outcome was presented. To ensure attention to the cues, participants indicated (at the time of the outcome) whether the outcome was

expected or not with a button press. I specifically investigated passively delivered rather than instrumentally obtained rewards so that any activation differences between self and other trials could not be related to differences in motor preparation (i.e. for example, an action on a self trial but no action on another trial).

Prior to scanning participants completed a practice version of the task during which they received feedback as to whether their judgements (expected or unexpected) were correct. During scanning, however, participants were instructed that they would not receive feedback on their judgements but that they should respond as quickly and accurately as possible to the judgement.

There were 100 trials in total, 50 self trials and 50 other trials presented in a pseudo-random order, with no more than 3 trials in a row of self or other cues. The 50 self trials consisted of 25 trials of high probability first-person cues and 25 trials of low probability first-person cues. Similarly, the 50 other trials consisted of 25 high probability third-person cues and 25 low probability first-person cues. For both self and other conditions, 20 outcomes were an expected win, 20 outcomes were an expected no win, 5 outcomes were an unexpected win and 5 outcomes were an unexpected no win (equivalent to 80%/20% probability).

5.2.2.2. Trial structure

Each trial began with a cue that signalled the probability of reward (80%/20%) and agent (self/other) for 800 ms (see Figure 1A). After a jittered delay (2500-6000 ms) participants observed an outcome (win 100 points/win 0 points) (800 ms) followed by a variable fixation (2000-4000 ms). Participants were then presented with the options ‘YES’ ‘NO’ where they were required to press one of two buttons to indicate whether the outcome was expected or not. The side of the screen that these options were presented on was counterbalanced so that participants could not form a representation of a specific motor command at any

point during a trial. Participants had 1500 ms to indicate their option, or the word 'MISSED' appeared in red on the screen. This was then followed by a fixation cross (1000-2000 ms).

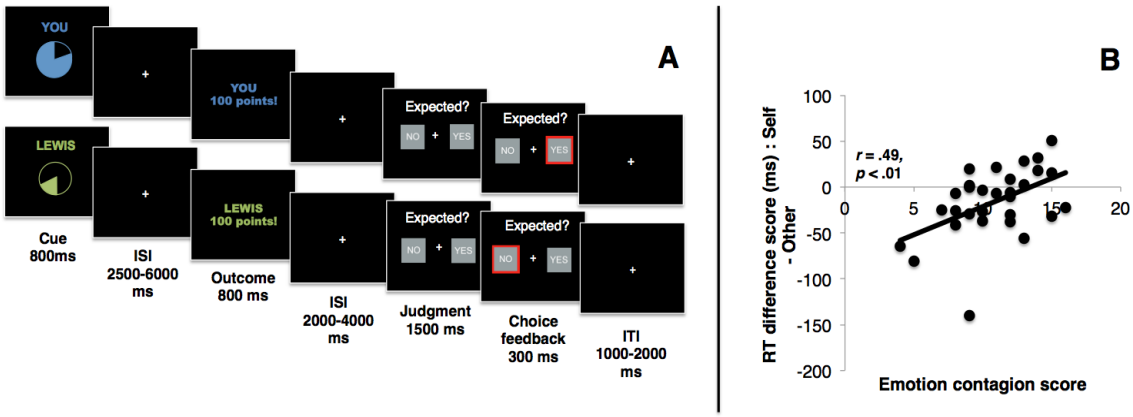
5.2.3. *Procedure*

Participants were paired with one of two age-matched confederates (who were also male), whom they believed were naïve participants and had never met prior to the experiment. The participant and the confederate were instructed together that they could earn extra payment, based on the outcomes they received during the experimental task (see below); but in fact all participants were paid the same amount (total £30, representing an additional £7 to the standard participant payment for the required time commitment). They also believed that the confederate participant could earn an extra payment in the same manner during the task. A set of standardised questions completed after the scan confirmed that no participant had become suspicious about the deception during the experiment.

Participants attended two sessions. The first session was attended only by the experimental participant without a confederate and involved practicing the experimental task and completing questionnaires. In the first session, attended only by the experimental participant, the "other" participant was described as 'Player 2' and the experimental participant was instructed that in the scanning session this name would be replaced by the name of the other participant. Participants were instructed that during the practice session the points would not be converted into any money either for themselves or the other person but that when they attended the scanning session these points would be converted into additional payment for themselves and the other participant. The second session (<7 days later) was attended by both the experimental participant and the confederate. During this session the experimental participant performed the task whilst inside the MRI scanner. The experimental participant was under the impression that the confederate performed the same task, simultaneously. The

confederate was seated in the adjacent MRI control room to maintain this impression. The participants were instructed that regardless of whether the cues and outcomes were for themselves or for the other person, that they should perform the same judgment task to indicate whether the outcome was expected or not. Moreover, participants were not instructed to the specific payoff matrix, which was in fact equal. This was done to ensure participants remained motivated to attend to the cues and outcomes.

Figure 5.1. (A) Trial structure for vicarious reward prediction task. (B) behavioural correlation.



Notes: (A) Participants performed trials that began with a cue signalling the probability of reward (high [80%] or low [20%]) and the agent to whom reward would be delivered (self - 'YOU'; or other - 'LEWIS' in this example). Participants judged whether the outcome (win 100 points or win 0 points) was expected or unexpected after outcome delivery. Participants believed that the other participant outside of the scanner was simultaneously performing the same task and that the points they observed would be converted into additional payment at the end of the experiment for themselves and for the other participant. **(B)** Scatterplot showing association between self-other RT difference at the time of the judgement and trait emotion contagion (n=30). Overall, participants were slower when making judgments about the expectedness of outcomes for other compared to self. However, this effect was associated with emotion contagion such that those highest in emotion contagion showed a relative speeding of response for other.

After the scanning session, participants rated how positive they felt when observing themselves or the other person winning on a 9 point scale ranging

from “not at all” to “very positive”. One-sample t-tests showed that for both self and other, participants felt significantly more positive than neutral when seeing win outcomes compared to no win outcomes (Other win $t(29)=2.1$, $p<.05$, $M=5.4$, $SD=.1.04$, Self win $t(29)=5.3$ $p<.001$, $M=6.4$, $SD=1.43$).

5.2.4. *Questionnaire measures*

Participants completed a measure of empathy, the Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011). As outlined in the introduction, the QCAE is a multidimensional instrument devised to measure five key components of empathy. In the development of the QCAE, two raters selected items from other well-validated and commonly used empathy measures if they were deemed to measure empathy (see items below). Items deemed to measure other processes (e.g. sympathy) were not included. The five subscales comprising the QCAE are: perspective-taking (e.g. ‘‘I can easily tell if someone else wants to enter a conversation.’’); online simulation (e.g. ‘‘Before criticizing somebody, I try to imagine how I would feel if I was in their place.’’); emotion contagion (e.g. ‘‘I am happy when I am with a cheerful group and sad when the others are glum.’’); peripheral responsivity (e.g. ‘‘I often get deeply involved with the feelings of a character in a film, play, or novel.’’); and proximal responsivity (e.g. ‘‘I often get emotionally involved with my friends’ problems’’). Items are rated on a four-point scale from ‘‘strongly disagree’’ to ‘‘strongly agree’’. The QCAE has good construct validity and internal consistency (Reniers et al., 2011).

5.2.5. *Statistical analysis of behavioural data*

Behavioural analyses were performed in SPSS 22 (Armonk, New York: IBM Corp). An agency (self vs. other) by reward (win vs. no win) analysis of variance

(ANOVA) was used to examine reaction time (RT) differences to outcome judgments. I did not examine the agency (self vs. other) by expectedness (expected vs. unexpected) interaction due to the low number of unexpected outcomes in this design (<10 valid trials per subject). Relationships between behavioural performance and empathy were assessed using bivariate correlations. I adopted an alpha level of 0.05, and a power analysis indicated that I had ~80% power to detect an effect size of Cohen's d=0.50.

5.3. Functional neuroimaging data collection and analysis

5.3.1. *fMRI data acquisition*

A Siemens Avanto 1.5-T MRI scanner was used to acquire a 5.5-minute 3-dimensional T1-weighted structural scan and 424 multislice T2*-weighted echo planar volumes with blood oxygenation-level–dependent (BOLD) contrast. The structural scan was acquired using a magnetization prepared rapid gradient echo (MPRAGE) sequence. Imaging parameters were: 176 slices; slice thickness=1 mm; gap between slices=0.5 mm; TR=2730 ms; TE=3.57 ms; field of view=256 mm x 256mm2; matrix size=256 x 256; voxel size=1×1×1 mm resolution. The EPI sequence was acquired in an ascending manner, at an oblique angle (≈30˚) to the AC-PC line to decrease the impact of susceptibility artefact in the orbitofrontal cortex (Bird & Viding, 2014; Blair, 2005)with the following acquisition parameters: 424 T2*-weighted echo planar volumes, 35 2mm slices, 1 mm slice gap; echo time=50 ms; repetition time=2975 ms; flip angle=90°; field of view=192 mm; matrix size=64x64.

5.3.2. *fMRI data analysis*

Imaging data were analysed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Data pre-processing followed a standard sequence. The first four volumes were discarded to allow for T1 equilibration effects and last volume was discarded as the experimental task ended one volume before the end of the scanning sequence. Images were then realigned and co-registered to the participant's own anatomical image. The anatomical image was processed using a unified segmentation procedure combining segmentation, bias correction, and spatial normalization to the Montreal Neurological Institute (MNI) template using SPM's New Segment procedure (Ashburner & Friston, 2005); the same normalization parameters were then used to normalize the EPI images. The images were resampled to a voxel size of 1.5 x 1.5 x 1.5 mm. Finally, a Gaussian kernel of 8 mm full-width at half-maximum was applied to spatially smooth the images. Before the study, first-level design matrices were examined to ensure that estimable GLMs could be performed with independence between all regressors, with correlation coefficients of $r < 0.25$.

5.3.3. *First-level analysis*

Nine (in some subjects ten) event types were used to construct regressors in which event onsets were convolved with the synthetic canonical haemodynamic response function in SPM, and associated responses were estimated in the context of the general linear model. Each of the four conditions (self high probability, self low probability, other high probability, other low probability) at the time of the cue and at the time of the outcome were modelled as separate regressors for correct responses. The onset of the judgement was also modelled, in a single regressor across all event types. An additional regressor modelled trials where the judgement was missed or participants made an error. For those participants where head motion caused visible image corruption in particular scans an extra regressor was included. These images were removed and replaced with an image created by interpolating the two adjacent images in order to prevent distortion of the between-subjects mask (four participants, in

each accounting for <1% of the total fMRI data). The residual effects of head motion were also modelled as covariates of no interest in the analysis by including the six head motion parameters estimated during realignment. Data were high-pass filtered at 128 s to remove low-frequency drifts, and the statistical model included an AR(1) autoregressive function to account for autocorrelations intrinsic to the fMRI time-series. Contrast images were computed to examine the interaction (agency x probability), and main effects of agency (self > other and other > self) and probability (high> low and low> high) at the time of the cue.

Many studies have suggested that situations which involve mixed pay-offs between study participants and other people can result in neural responses that reflect payoff differences between self and other, i.e. they relate to coding of rewards for self relative for other, often called 'inequity aversion' rather than 'vicarious' reward responses (see Ruff & Fehr, 2014 and Rilling & Sanfey, 2011 for reviews). To examine whether identified neural responses in the current study to reward predicting cues were reflective of coding of rewards for self relative to other, and thus inequity aversion, I constructed a second model that was the same as the main model but contained all cues collapsed into a single regressor. This regressor had two associated parametric modulators. The first coded the "inequity" - the difference in accumulated reward between self and other on each trial - and the second coded the agent x probability interaction. This allowed us to examine (1) neural responses to inequity and (2) whether any neural responses occurred over and above the variance explained by inequity.

5.3.4. *Second-level analysis*

Second-level analysis was performed using the standard summary statistics approach to random effects analysis in SPM. Contrast images were input into second-level one-sample t-test design matrices. Interactions and main effects are reported at $p < .05$, family-wise error (FWE) corrected at the voxel level

across the whole brain. Where significant interactions were identified, I conducted illustrative post-hoc analyses with simple main effects contrasts using a less conservative statistical threshold of p < .001 (uncorrected).

5.4. Results

5.4.1. *Behavioural data*

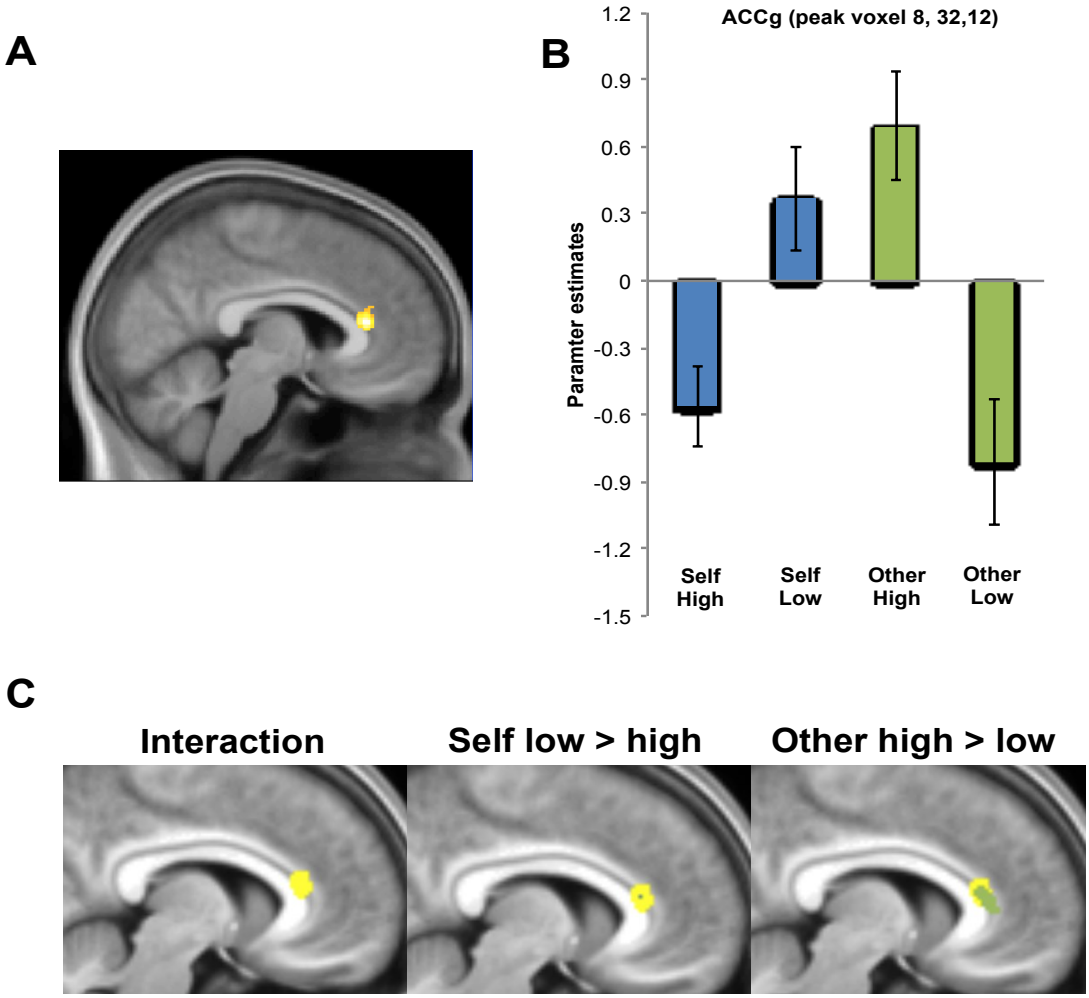
Participants were highly accurate in their judgments of whether the outcome was expected or not (mean accuracy >91% for all participants for both trial types) and missed very few trials (mean <1% for all participants). For mean RTs, a 2 (self vs. other) by 2 (win vs. no win) ANOVA showed significantly slower judgments on third-person (M=664 ms, SD=18) than on first-person (M=649 ms, SD=16) trials (main effect of agency: F (1,29) = 5.32, p = .03). Judgments were also significantly faster after reward (641ms, SD = 16) compared to a no reward (672ms, SD=19) outcomes (main effect of outcome F (1,29) = 14.34, p = .001). The agency x reward interaction was non-significant (F (1,29) = .05, p = .83).

Given the significant main effect of agency, I calculated the difference score between self and other RTs to examine associations between this behavioural measure and empathy. The emotion contagion subscale of the QCAE was positively associated with the self–other difference score (r = .49, p < .01), i.e. participants higher in emotion contagion showed a relative facilitation (speeding) when making decisions about the expectedness of outcomes for other people (**Figure 1, Panel B**). No other subscale of the QCAE correlated with the self-other difference score (all ps > .49). Multiple regression including all QCAE subscales showed that the association between the self-other difference score and self-reported empathy was specific to the emotion contagion subscale (β = .55, SEM = 2.43, p < .01).

5.4.2. *fMRI data: agency x probability interaction at time of the cue*

To test the first hypothesis, that activity in the ACCg would signal information about reward probability for others, I examined the agency x probability interaction at the time of the cue. In line with this hypothesis, this analysis revealed a significant effect in the ACCg (MNI coordinates [x=8, y=32, z=12], Z = 5.05, k=10, $p < .05$ FWE, whole brain corrected), putatively in area 24a'/24b' at the border of the midcingulate and anterior cingulate sub-regions (**Figure 5.2**). I examined the nature of this interaction by testing the simple main effects, specifically the contrasts of other high vs. low probability and self low vs. high probability. Inspection of the other high vs. low probability simple main effect revealed a large cluster in the ACCg overlapping with the region identified in the interaction (MNI coordinates [x=6, y=33, z=12], Z = 4.14, k=184, $p < .001$ (uncorrected)). Inspection of the self low vs. high probability contrast revealed a small cluster of overlapping voxels (MNI coordinates [x=9, y=32, z=13], Z = 3.28, k=5, $p < .001$ (uncorrected)). This suggests that the ACCg activation identified in the interaction mainly signals the probability of rewards that would be received by another person.

Figure 5.2. Interactions and main effects in ACCg.



Notes: Activation in the ACCg (A, C-left) signaled the agency (self vs. other) by probability (high [80%] or low [20%]) interaction at the time of the cue [x=8, y=32, z=12], displayed at p <. 001 (uncorrected). (B) Parameter estimates for the peak voxel in the ACCg. (C) Left: overlay of the agency x probability interaction in ACCg (yellow, as in A). Middle: only a small number of voxels overlapped between the interaction contrast (yellow) and the simple main effect of self low vs. high probability (blue, k=5 at p<.001 uncorrected). Right: a large number of voxels overlapped between the interaction contrast (yellow) and the simple main effect of other high>low probability (green, k=184 at p<.001 uncorrected). Error bars indicate SEM.

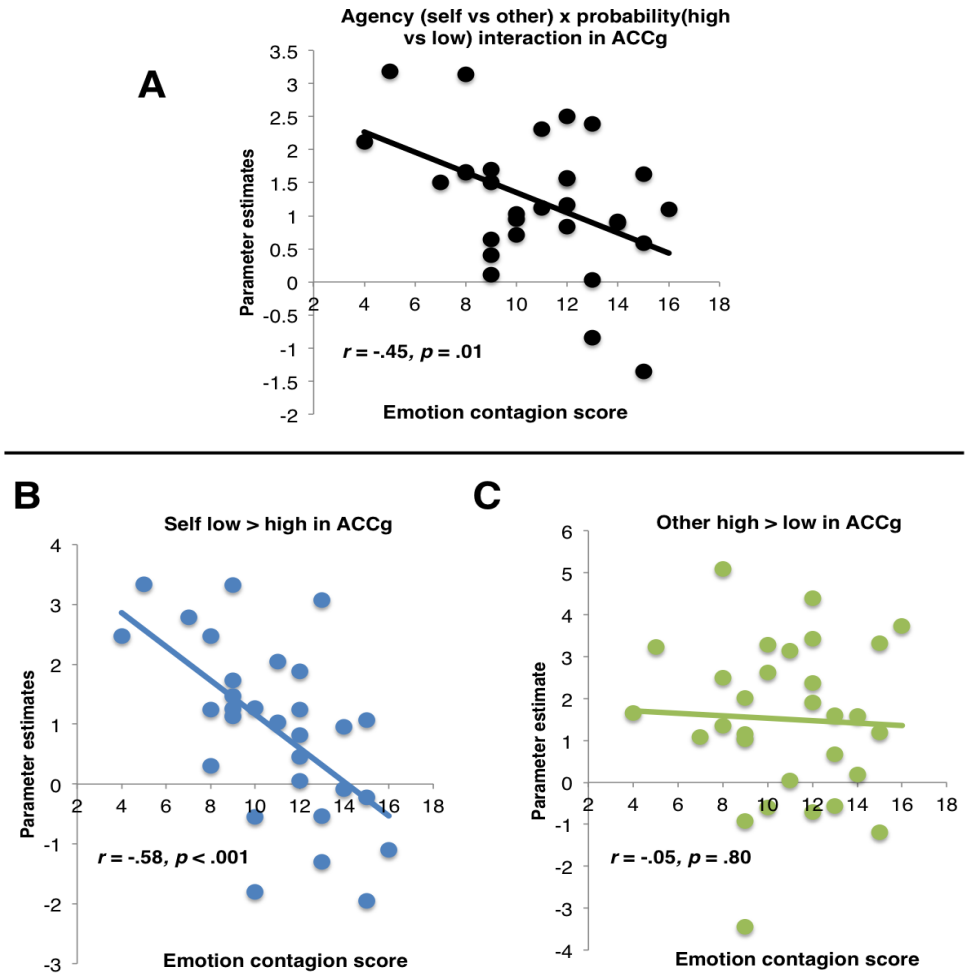
5.4.3. *fMRI data: associations with trait empathy*

To test the second hypothesis, that the extent to which ACCg responds to the probability of rewards specifically for others would be positively associated with trait empathy, I used MarsBaR (Brett, Anton, Valabregue, & Poline, 2002) to extract individual interaction contrast estimates (other high vs. low probability minus self high vs. low probability) from the ACCg cluster identified above, and correlated these with participants' self-reported empathy on the five QCAE subscales. Emotion contagion was significantly negatively associated with the ACCg interaction contrast estimate (r = -.45, p = .01, all other subscales p > .58) and multiple regression including all QCAE subscales showed that this effect was specific to emotion contagion (β = -.60, SEM = .062, p = .003, all other subscales *ps*>.15) (**Figure 5.3**). In other words, the interaction was weakest in individuals high in emotion contagion.

To better understand the nature of this association, I examined the correlations for other high vs. low probability and self low vs. high probability in ACCg with empathy subscales (see **Figure 5.3**). There was no significant correlation between ACCg response to other high vs. low probability (r = -.05, p = .81) and empathy. However, there was a significant negative association between ACCg response to self low vs. high probability and emotion contagion (r = -.58, p < .001) and these correlations were significantly different from one another (Z=-2.0, p<.05); again, multiple regression demonstrated that this effect was unique to emotion contagion (β = -.66 SEM = .082, p < .001, all other subscales *ps* > .19). In other words, the extent to which ACCg distinguished between low and high reward probability for self was attenuated in individuals with high emotion contagion.

In summary, in individuals with high emotion contagion, the ACCg signalled information about the relative difference between high and low probability rewards only for others; whilst in individuals with low emotion contagion the ACCg additionally signalled (negatively) reward probability for self.

Figure 5.3. Correlations with significant interaction effect in ACCg.



Notes: (A) Significant association between the cluster in the ACCg showing the interaction effect and participants' emotion contagion scores. **(B)** Response to self low > high probability decreases as a function of emotion contagion, with those lowest in emotion contagion showing the greatest response to low>high probability of reward for self. **(C)** Response to other high > low probability shows no significant modulation as a function of emotion contagion.

5.4.4. *fMRI data: main effects at the time of the cue*

The temporal pole showed a significant main effect of other > self (MNI coordinates [33, 22, -26]; z = 4. 85; k = 2, $p < .05$, FWE whole brain corrected).

No other main effects or interactions survived whole-brain correction for multiple comparisons. I provide uncorrected tables thresholded at $p < .001$ k=10 for completeness (**Appendix 3**). I note that these results should be interpreted with caution given that they do not survive correction for multiple comparisons.

5.4.5. *fMRI data: agency x outcome interaction and main effects at the time of the outcome*

No interactions or main-effects survived whole brain-correction for multiple comparisons. I provide uncorrected tables thresholded at $p < .001$ $k = 10$ (**Appendix 4**) for completeness. I note that these results should be interpreted with caution given that they do not survive correction for multiple comparisons.

5.4.6. *fMRI data: analysis of inequity aversion*

Analysis of the inequity parametric modulator showed no whole brain corrected results and no uncorrected results in ACCg. I then tested whether the observed effects in the ACCg occurred over and above any effects of inequity. This analysis showed that there was still a significant effect in the ACCg after accounting for the variance explained by inequity (MNI coordinates [x=6, y=32, z=13], Z= 4.97, k = 8, p < .05 FWE, whole brain corrected). Consequently, the ACCg response was unlikely to reflect differences in accumulated reward between self and other.

5.5. Discussion

I examined haemodynamic responses in the human brain to cues that predicted a high or low probability of a reward for oneself or another person. I show that the ACCg robustly signals the probability of rewards for another person. This supports the hypothesis that the ACCg is engaged when processing predictions about rewards for other people. My second hypothesis that that the extent to which the ACCg is specialised for processing others' rewards is positively associated with trait empathy was partially supported. As predicted the interaction effect in the ACCg significantly co-varied with emotion contagion. However, this effect was driven by the extent to which ACCg signalled reward predictions for self, not other. Specifically, for those high in emotion contagion the ACCg signalled reward prediction exclusively for others, whilst for those low in emotion contagion this same region signalled reward prediction for self (in the opposite direction).

The model of the contributions of ACCg to social cognition (Apps, Lockwood, et al., 2013) highlights that this region plays an important role in understanding the value of others' rewards, and consequently in social behaviour (Apps, Green, et al., 2013; Apps & Ramnani, 2014; Boorman et al., 2013; Chang et al., 2013; Jones et al., 2011; Rudebeck et al., 2006). This claim is built upon several lines of evidence. Lesions to this region have been shown to impair the processing of social stimuli and cause a reduction in the execution of social behaviours (Rudebeck et al., 2006). The ACCg is connected to regions that process social information, but also, to regions that process reward-related information (Haber et al., 1995; Lynd-Balta & Haber, 1994; Yeterian & Pandya, 1991). Single-unit recording evidence suggests that a relatively large proportion of ACCg neurons, compared to those in other prefrontal regions, respond when a monkey anticipates the delivery of reward to another monkey (Chang et al., 2013), and human imaging studies have shown that the ACCg responds when tracking the value of cues predicting approval from peers (Jones et al., 2011). Taken together these studies support the claim that the ACCg is important for processing others' rewards and also in social behaviour. However, a key

untested component of this model was that the ACCg would be engaged when processing the likelihood of rewards being delivered to others. I show for the first time that the ACCg signals the likelihood of others' rewards, regardless of trait levels empathy. I also note that I did not observe responses to reward prediction in other candidate regions for reward signals, even at uncorrected levels (e.g. ventral striatum, ventromedial prefrontal cortex and amygdala, see Morelli et al., 2015 for a meta-analysis), supporting some degree of specificity of ACCg response to vicarious rewards in this study.

The experimental paradigm was designed to ensure participants attended to reward cues. By asking participants to make a decision at the time of the outcome I cannot purely assess whether outcome related responses are also coded in ACCg, as participants were both processing the outcome and preparing a motor response during this time. However, there is evidence that vicarious prediction error signals may well be coded in ACCg (e.g. Apps et al., 2013; 2015). I provide the first evidence that this same region also encodes the likelihood of others receiving rewards.

Whilst previous studies have suggested the ACCg plays an important role in empathy (Lamm et al., 2011; Engen & Singer, 2013) these studies have largely focused on neural responses to others' pain. These data suggest that the degree of specialisation in this region's response to others' predicted rewards may partly underlie individual differences in emotion contagion. Emotion contagion is hypothesised to be a necessary foundation for empathising with other individuals (e.g. Bird and Viding, 2014) and is a process that is shared with non-human animals (reviewed in (De Waal, 2008). Importantly, emotion contagion also covaried with RTs to decisions about rewards delivered to others, with those highest in trait emotion contagion showing the greatest speeding of response. A distinction is often made between 'affective empathy', commonly understood as an affective state caused by vicariously processing the experiences of another person, and 'cognitive empathy' which is thought to include processes such as perspective-taking and theory of mind (Singer & Lamm, 2009). Regression analyses suggested that only emotion contagion, part of the 'affective' component, was associated with vicarious reward prediction. In tasks investigating cognitive aspects of empathy an anatomically

separate region of the medial prefrontal cortex (mPFC), the dorsal mPFC is often responsive (e.g. Amodio & Frith, 2006; Frith & Frith, 2006), suggesting partially separate functions of ACCg and mPFC.

Although I did not predict an association between emotion contagion and ACCg response to self reward prediction, a possible explanation relates to the findings of Chang et al., (2013) and Haroush and Williams (2015). These authors observed some 'self-reward' selective neurons in the same region of the ACCg/dACC that also contained 'other-reward' selective neurons, suggesting that some processing of information about rewards for self occurs in ACCg. However, given the limited sample sizes in non-human primate studies these authors were unable to examine variability in the proportion of neurons that signalled self vs. other reward. I speculate that even if at the population level the ACCg shows a relative specialisation in processing rewards for others, individual variability in the degree to which self rewards are also processed in this region could be important for explaining heterogeneity in ACCg function and empathy. That is, for those individuals who display the lowest levels of emotion contagion, there appears to be reduced specialisation and a potentially opposing coding scheme of self and other reward probability in ACCg. Such opposing coding within the same anatomical region could have consequences for understanding social cognition and behaviour, such as increased weighting of rewards to self and higher likelihood of engaging in competitive social interactions.

This interpretation is supported by a recent study, which found that stimulation of dACC neurons made monkeys more competitive (Haroush and Williams, 2015). Similarly, another study showed that single neurons in a region of the rat cingulate cortex thought to be homologous with human dACC, coded the value of competing with another rat for rewards (Hillman & Bilkey, 2012). These findings may help reconcile previous discrepancies in the functions that have been imputed to dACC in terms of competitive social behaviours (Hillman & Bilkey, 2012; Haroush & Williams, 2015) but also empathy (Lamm et al., 2011; Engen & Singer, 2013). I propose that variability in empathy may modulate not only the extent to which social information is processed in ACCg, as suggested in previous studies and theoretical accounts of empathy (e.g. Lamm et al.,

2011; Engen & Singer, 2013), but also the extent to which self as well as other reward information is computed. However, this hypothesis requires further testing in future experiments.

Empathic abilities are a fundamental building block for successful social behaviour and are at the core of many disorders of social cognition, including autism and psychopathy (Blair, 2005; Bird & Viding, 2014; **Chapter 2**). Previous studies have suggested that a similar portion of the dACC that was activated in this study is anatomically and functionally atypical in individuals with psychopathy and in individuals with autism (e.g. (Anderson & Kiehl, 2012; Brazil et al., 2011; Delmonte, Gallagher, O'Hanlon, McGrath, & Balsters, 2013; Simms, Kemper, Timbie, Bauman, & Blatt, 2009).

Integrating these previous findings with the present results suggests the hypothesis that individual differences in the structure, function and connectivity of the ACCg constrain the extent to which this region processes reward predicting cues for others compared to self, which may lead to atypical empathic processing. However, individuals with psychopathy and autism have different profiles of empathic processing and behaviour from one another (Blair, 2005; Bird and Viding, 2014; **Chapter 2**). The ACCg has strong connections to other regions involved in social and reward processing including the nucleus accumbens (Anderson & Kiehl, 2012; Brazil et al., 2011; Delmonte, Gallagher, O'Hanlon, McGrath, & Balsters, 2013; Simms, Kemper, Timbie, Bauman, & Blatt, 2009), a region also suggested to participate in vicarious reward processing (Braams et al., 2014; Fareri, Niznikiewicz, Lee, & Delgado, 2012; Mobbs et al., 2009), the temporal poles (which showed greater response to other vs. self reward prediction in this study), the temporo-parietal junction and paracingulate cortex (Barbas, Ghashghaei, Dombrowski, & Rempel-Clower, 1999; Markowitsch, Emmans, Irle, Streicher, & Preilowski, 1985; Seltzer & Pandya, 1989). Future research into the neurocognitive correlates of psychopathy and autism should examine whether distinct social behavioural abnormalities can be characterised by differences in the functional and connectional fingerprint of the ACCg during vicarious reward processing.

5.6. Conclusions

In summary, I demonstrate a central role for the ACCg in processing predictions about the likelihood of others' rewards. I also found substantial individual variation in the degree to which the ACCg responds to self and other reward, with only those highest in trait emotion contagion showing specialisation of ACCg for others predicted reward. Taken together, these findings highlight the importance of understanding the contributions of the ACCg to social cognition and how variability in its function may underlie variability in social behaviour.

In the final empirical chapter, I present a study that examines vicarious decision-making, and combines computational modelling of behaviour, an assessment of trait empathy and fMRI to examine the neural and behavioural mechanisms that link empathy to prosocial behaviour.

CHAPTER 6: Chapter background

In the last chapter, **Chapter 5**, I examined neural responses to vicarious reward and their modulation by individual differences in empathy. In the current chapter I aim to identify the behavioural and neural mechanisms that could link empathy to prosocial behaviour.

Representing and vicariously processing rewards that others receive following our choices is likely to be of central importance when humans learn to be prosocial. In reinforcement learning (RL) theory, a prediction error (PEs) – the difference between and expected and actual outcome as a choice – is the key computation that drives learning. Many studies have highlighted RL mechanisms as crucial for learning across species. We can apply this same framework to understand how we make prosocial choices, where we learn about rewards that others receive following our choices.

In this Chapter I used RL theory as a model to try to understand prosocial behaviour and combined computational modelling of behaviour with neuroimaging and a trait measure of empathy. Participants (n=31) performed a reinforcement learning based task in which they were required to learn the probability that each of two stimuli would be rewarded (high probability vs. low probability). They performed this task for themselves (self reinforcement condition), for another participant (confederate, prosocial reinforcement condition) or for no one (no reinforcement, control condition). Using a RL algorithm, I was able to model differences in the learning rates between these conditions and neural response to PEs in key neural regions associated with social decisions and/or social PEs in previous studies, ventral striatum, sgACC, ACCg, OFC and DLPFC. Individuals higher in empathy may be more likely to engage in prosocial behaviours (Eisenberg, Eggum, & Di Giunta, 2010; Hoffman, 2008; **Chapter 3**). Thus, I also examined how individual differences in empathy modulated prosocial learning behaviour and neural response to identify mechanisms linking empathy to prosocial behaviour.

6.1. Introduction

Humans have a remarkable capacity to engage in prosocial behaviours, even with genetically unrelated individuals (Fehr & Fischbacher, 2003, 2004). People routinely participate in charitable donation and exhibit social preferences, which are influenced by concern for the welfare of others (Fehr & Camerer, 2007). Empathy, the capacity to vicariously experience and to understand the affect of other people (Bird & Viding, 2014; Decety & Jackson, 2004; Eisenberg, 2000; Hoffman, 2001; Singer & Lamm, 2009) is suggested to be an important facilitator of such behaviours (Eisenberg, Eggum, & Di Giunta, 2010; Hoffman, 2008; **Chapter 3**), yet the mechanisms linking empathy to prosocial behaviour are not yet fully understood.

The behavioural and neural basis of prosocial behaviours have often been investigated in the context of economic decision-making tasks, such as the ultimatum game and dictator game, in tasks of moral decision-making where participants read and judge the permissibility of moral scenarios, or where they are asked to make decisions to donate to charity (see Rilling & Sanfey, 2011, Ruff & Fehr, 2014 and Moll & Schulikin, 2009 for recent reviews). In general, these paradigms have shown that a distributed set of neural regions, including the ventral striatum, subgenual cingulate cortex (sgACC), orbitofrontal cortex (OFC) and dorsolateral prefrontal cortex (DLPFC) respond in the context of making social decisions (Rilling & Sanfey, 2011; Ruff & Fehr, 2014; Moll & Schulikin, 2009).

A separate line of research has applied reinforcement learning models of reward guided behaviour, to understand decision-making motivated by self and social preferences. In these models, the prediction error (PE) – the difference between a predicted and actual outcome of a choice – act as a key signal to drive learning (Sutton & Barto, 1998). Many animal studies have identified neurons that signal PEs when the outcomes of one's own actions have unexpected consequences for oneself (Rushworth, Mars, & Summerfield, 2009). However, we can apply this same framework to understand prosocial

learning behaviour – where we makes decisions that have consequences for another person and learn from outcomes that others receive following our choices.

Using RL theory as a model to understand prosocial behaviour provides an important compliment to economic games and moral scenarios. Such models contain parameters that track people's learning over time in both social and nonsocial contexts, which can be sensitive in predicting people's future choices. In functional magnetic resonance imaging (fMRI), these models allow us to examine neural responses parametrically rather than relying on subtraction based designs, which average over trials. Hypotheses can be formulated not only about the function of specific brain regions, but also about the computational mechanisms they may underpin. Finally RL theory has been shown to characterise learning in many different contexts and across different species (Behrens et al., 2009).

In one of the first paradigms to examine self and social PE signalling, prediction errors were evoked by unexpectedly changing the outcome of a decision to donate to charity to be either for the participant themselves or for the charity (Harbaugh et al., 2007). The authors found PE signals in the ventral striatum that responded to rewards for the participant themselves but also for a charity (Harbaugh et al., 2007). Recently, it was shown the ventral striatum signalled PEs when learning to obtain points to reduce the amount of noise another participant will experience, as well as noise to be experienced by oneself (Sul et al., 2015). Social PE and reward signals have also been found to covary with learning about reward outcomes by observing another person's actions and their outcomes, identified in DLPFC and VMPFC (Burke et al., 2013), in the gyral portion of the ACC (ACCg) when learning about the volatility of social advice (Behrens et al., 2008) and in the sgACC when making decisions to donate to charity vs. oneself (Moll et al., 2006).

A key question that remains unanswered is whether these regions involved in decision-making contexts that involve other people (what Ruff & Fehr (2014) call 'socially-specific' decision making) are the same as those involved in

decision-making for self (what Ruff & Fehr (2014) call 'non-social' decision making) – i.e. is there 'common currency' of PEs that supports social and non-social decision learning (Ruff & Fehr, 2014). There is support for both positions, with regions such as the ventral striatum observed to respond to both social and non-social PEs (e.g. Harbaugh et al., 2007; Sul et al., 2015) whilst regions such as the ACCg and sgACC appear to code for others' rewards exclusively (e.g. (Apps, Green, et al., 2013; Apps et al., 2015; Behrens et al., 2008; Moll et al., 2006). In contrast, regions such as the OFC and the DLPFC respond in tasks whether there is a conflict between self and social preferences (e.g. (Lee, 2008; Sanfey, 2007), which suggests that these regions could differentially process self and prosocial reward. However, the tasks that have been used to study 'socially-specific' and 'non-social' decision making have not included a non-self reward control condition. This means that it is currently unclear whether 'social' PEs relate to social computations exclusively or reflect foregone/fictive rewards, i.e. rewards that are not delivered to oneself (Hayden et al., 2009; Lohrenz et al., 2007). Moreover, we know that there are large individual differences in people's motivation for prosocial behaviour, with some individuals motivated by selfish preferences and some by prosocial preferences (Sul et al., 2015). Such individual differences may be, in part, underpinned by variability in empathy (Eisenberg, Eggum, & Di Giunta, 2010; Hoffman, 2008; **Chapter 3**). Yet the mechanisms by which trait individual differences in empathy relate to differences in prosocial learning behaviour and neural response are still to be identified.

In this study I used reinforcement learning theory as a model for understanding prosocial behaviour. I designed a novel prosocial reinforcement-learning (PRL) task where participants chose between one of two stimuli on each trial that were probabilistically associated with a reward outcome. One stimulus was always associated with a high probability (75%) of receiving a reward and the other with a low probability (25%). Participants were instructed that they would do the task under three different conditions where all outcomes would be received 1) by them (self reinforcement condition), 2) by another participant (confederate, prosocial reinforcement condition), or 3) by no one (no reinforcement, control condition). Using a RL algorithm, I was able to examine differences in the

learning rates between these conditions and neural response to PEs in key neural regions associated with social decisions and/or social PEs in previous studies, ventral striatum, sgACC, ACCg, OFC and DLPFC. There is evidence that empathy can modulate processing of social information, such that those who are higher in empathy may also be more likely to engage in prosocial behaviours (Eisenberg, Eggum, & Di Giunta, 2010; Hoffman, 2008; **Chapter 3**). Thus, I also examined how individual differences in empathy modulated prosocial learning behaviour and neural response to identify the mechanisms that link empathy to prosocial behaviour.

6.2. Method

6.2.1. *Participants*

Thirty-four right-handed healthy males (age 19-32, M=22.7 SD=3.0) were recruited through university participant databases. Exclusion criteria included previous or current neurological or psychiatric disorder, non-normal or non-corrected to normal vision, non-native English language and previous or current study of psychology. This latter criterion was employed due to concerns that prior experience of studying psychology could compromise participants' belief in the deception used in the protocol. Three participants were excluded from the analysis (two due to performance below 50% in all learning conditions and one due to neurological abnormalities) leaving a final sample of 31. All participants gave written informed consent and the study was approved by the local departmental research ethics committee.

6.2.2. *Procedure*

Participants were paired with one of two age-matched confederates (who were also male), whom they believed were naïve participants and had never met

prior to the experiment. Participants attended two sessions. The first session was attended only by the experimental participant and involved practicing the experimental task and completing questionnaires. The second session (< 7 days later) was attended by both the experimental participant and the confederate. The participant and confederate were taken together to the MRI centre and filled in consent forms together in the same room. The confederate was then led into a behavioural testing room and instructed to complete some questionnaires. The experimental participant was taken to the scanning room and reminded of the instructions for the task. They were told that they would view a pair of symbols on each trial and that they should select one of them. They would receive points for some of their choices that would be converted into money at the end of the experiment, such that the more points they received the more extra money they would earn. They were instructed that the two symbols would not be the same in terms of how often they gave points and with some symbols they were more likely to win points than other symbols. Whether the symbols appeared on the left or right did not affect their meaning. Finally, they were told that when they were playing for themselves they would receive any money they win.

Crucially, *when they were playing for the confederate, that participant would receive the money*. When they were playing for no one the points they saw would not be converted into any additional payment either for themselves or the other participant. Importantly, they were told that the other participant was not aware that they were performing a task where they could earn extra money and that any money they won would be given to the other participant anonymously, that is, it would be placed in a sealed envelope and the two participants would leave the scanning centre at different times.

Participants were instructed that they would receive extra payment based on the outcomes they received during the experimental task (see below); but in fact all participants were paid the same amount (total £30, representing an additional £7 to the standard participant payment for the required time commitment). They also believed that the confederate participant could earn an extra payment based on the choices the experimental participant made during

the task. A set of standardised questions completed after the scan confirmed that no participant had become suspicious about the deception during the experiment.

6.2.3. *Experimental task*

6.2.3.1. Design

The aim of this experiment was to examine BOLD signal that scaled parametrically with the size of a PE at the time of an outcome delivered to self, other or no one. Participants performed a probabilistic reinforcement-learning task where they were required to learn the probability that each of two symbols would be rewarded. One symbol of each pair was associated with a high probability (75%) and one with a low probability (25%) of reward. 'Self' blocks began with the instruction 'Play for YOU' and had the word 'YOU' written above all choice symbols and outcomes. 'Prosocial' blocks had the name of the other participant written above them (the names of the confederate participants). No one blocks had the word 'NO ONE' written above elements in a trial. This ensured that participants were explicitly aware whether the decisions they made resulted in outcomes for themselves, for the other participant or for no one.

Participants practiced one block (16 trials) of the task in a separate session ~7 days before the scanning session to familiarise themselves with the experimental task. During this block they were instructed that the outcomes would not be converted into any payment.

6.2.3.2. Trial structure

The beginning of each block began with an instruction screen that indicated the agent the outcomes would be received by (self, other participant, no one) for 2000ms (see **Figure 6.1**). This was followed by the presentation of two abstract

stimuli for 3000ms during which participants were required to select one of these. These stimuli were letters from the Agathodaimon font as in (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). If no response was indicated during this time the words 'MISSED' appeared in red on the screen. The option participants selected was shown for 300ms, followed by a delay (2500ms) then by the outcome of their choice (win 100 points/win 0 points). A variable fixation (2000-4000ms) was shown after the outcome before the two symbols were presented again (see Figure 1, Panel A). The side of the screen that the two symbols were presented was counterbalanced so that participants could not perform action-based learning.

There were 144 trials in total, 48 self, 48 other and 48 no one trials presented in three blocks of 16 trials. Each block began with a new pair of symbols to learn. Blocks were presented in pseudo-random order, with the same block type never presented twice in a row.

6.2.3.3. Questionnaire measures

Participants completed a measure of empathy, the Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011). The five subscales comprising the QCAE are: perspective-taking (e.g. ''I can easily tell if someone else wants to enter a conversation.''); online simulation (e.g. ''Before criticizing somebody, I try to imagine how I would feel if I was in their place.''); emotion contagion (e.g. ''I am happy when I am with a cheerful group and sad when the others are glum.''); peripheral responsivity (e.g. ''I often get deeply involved with the feelings of a character in a film, play, or novel.''); and proximal responsivity (e.g. ''I often get emotionally involved with my friends' problems''). Items are rated on a four-point scale from ''strongly disagree'' to ''strongly agree''. The QCAE has good construct validity and internal consistency (Reniers et al., 2011).

6.2.3.4. Computational modelling of behavioural data.

Learning behaviour in the self, other and no one conditions was modelled using a standard Rescorla-Wagner (R-W)-based reinforcement learning algorithm (Rescorla & Wagner, 1972), which has been extensively used to examine the behavioural and neural basis of arbitrary visuomotor associations in both self and social contexts (Brovelli, Laksiri, Nazarian, Meunier, & Boussaoud, 2008; Burke et al., 2010; Dayan & Balleine, 2002; Dayan & Daw, 2008; Schultz, 2006). The R-W model assumes that the associative value of an action (or stimulus) changes once new information reveals that the actual outcome of a decision is different from the predicted outcome (Rescorla & Wagner, 1972). Thus, on each trial, an action has a predicted associative value that is updated by a PE signal when the outcome reveals that this prediction is erroneous. At their most simple, RL algorithms state that expectations of future reward (Qt+1) should be a function of current expectations (Qt) and their discrepancy from the actual outcome that is experienced—the prediction error (δ). These reward predictions are updated by the learning rate (α):

(1) RL model

$$Q_{t(n+1)} = Q_{t(n)} + \alpha \, x \, \delta$$

Where:

(2) *Prediction error* $\delta = r_t - Q_{t(n)}$

In these equations *n* is the number of times an action, *t*, has been performed and *α* is the learning rate, in other words the extent to which the values are updated by new information. In (1) the value of the action in the future ($Q_{t(n+1)}$) is a function of current predicted value of the action (Q_t) added to the prediction

error (δ), which is multiplied by the learning rate (α). The learning rate defines the extent to which the prediction error updates the predicted value. Consequently, a low learning rate will minimise the influence of the prediction error and the amount that the value is updated. The prediction error, shown in (2), compares the actual outcome achieved by an action (r) to the prediction of its value ($Q_{t(n)}$). This difference is what determines the updating of the predicted value in the future.

To fit the R-W model to participants' behaviour I used the maximum a posteriori (MAP) approach (Daw, 2011). This is a two-stage procedure which begins by using a softmax function to estimate the probability of the subject choosing what they chose through maximum likelihood estimation. In the first stage, the softmax takes the predicted value from the Rescorla-Wagner model, and estimates a trial-by-trial probability of the subject choosing what they did given the model parameters (see equation 3). Within the softmax, there is an additional free parameter, called the inverse temperature, β, (sometimes referred to as the exploration/exploitation parameter) which estimates how stable the participant's choices are. The temperature parameter quantifies the noisiness of participants' choice behaviour, with a high parameter meaning very variable choice behaviour and a low parameter meaning very consistent choice behaviour.

(3) Softmax function

$$Pt(r) \quad = \exp\left(\frac{qt(r)}{\beta}\right) / \sum_n i = 1 \exp\left(\frac{qt(i)}{\beta}\right)$$

In order to estimate the learning rate and temperature parameters, every possible different combination of learning rates and temperature parameters was fitted to the data using the fmincon parameter search function in MATLAB. The probabilities output by the softmax were log-transformed and the parameters that "best-fit" (have the log-likelihood closest to 0) were selected.

In the second stage the group mean and standard deviation of the learning rate and temperature parameter was calculated to create a normal distribution around these values, which was then used as a prior for the learning rate and temperature parameters to be fitted to each participants choices a second time. This method provides a better estimation of each individual's true learning rate and temperature parameter that is less susceptible to the influence of outliers (Daw, 2011).

6.2.3.5. Model comparison

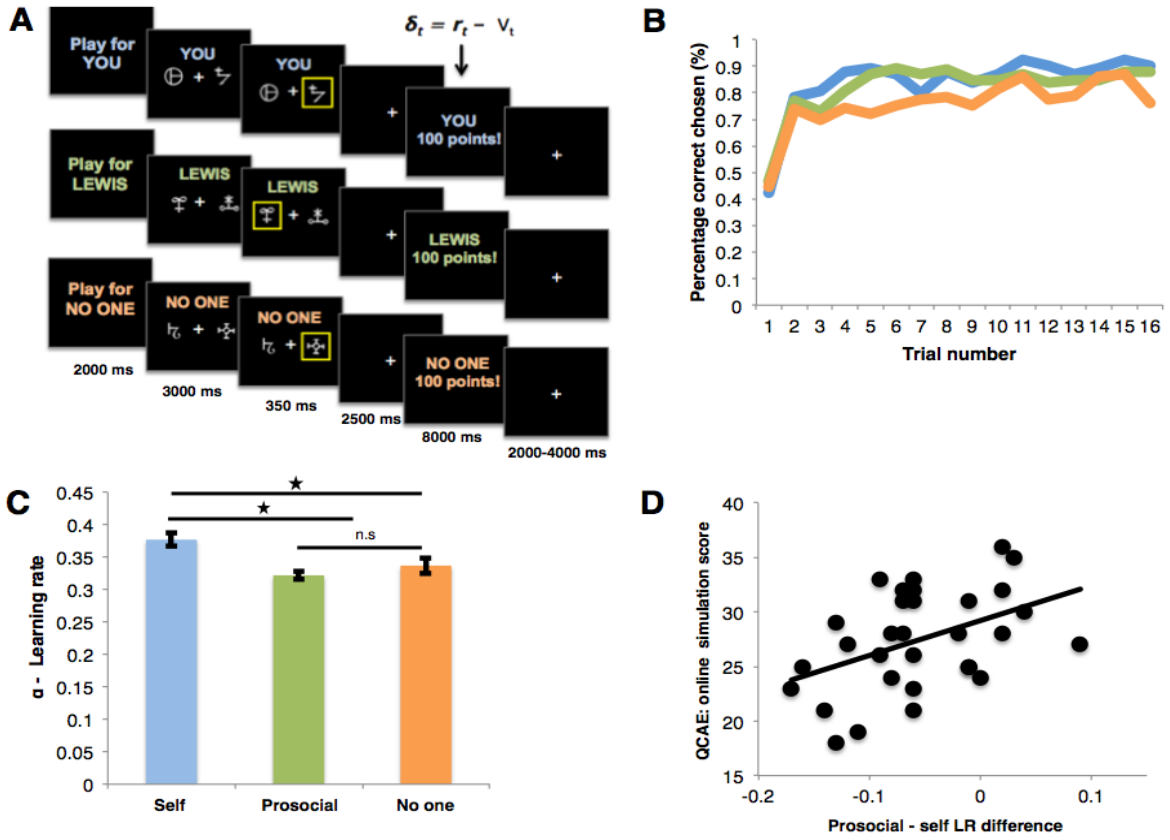
I compared the RL model to a null model where it is assumed that participants exhibit no learning and choose options at random. In this model the value for α was set to 0 to show that participants exhibited no learning and the value for β was varied between 0 and infinity. This null model was compared to the RL model using Bayesian Information Criterion scores (BIC) to examine whether participant's behaviour was better explained by an RL model compared to a model that assumed participants selected options randomly. The BIC compares the fit of the two models based on the number of parameters and the likelihood of the model fits, with models with fewer parameters being favoured. For all participants, their behaviour showed a better fit to the RL model compared to with the null model. This suggests that all participants learned during the task and a model that represents RL learning fits the data better than the null model, which represents random choices.

6.2.3.6. Statistical analysis of behavioural data

Behavioural analyses were performed in SPSS 22 (Armonk, New York: IBM Corp). We examined differences in both the learning rate and temperature parameters at the group level using a repeated measures ANOVA with there

levels (self, other and no one for both learning rate and temperature parameters). We also examined bivariate associations between the self-other learning rate and temperature difference and empathy components.

Figure 6.1. Trial structure and behavioural results.



Notes: (**A)** trial structure. Participants performed a reinforcement-learning task in which they had to learn the probability that abstract symbols were rewarded. One was always associated with a high probability and one with a low probability. At the beginning of the block participants were told whom they were playing for, either themselves, the other participant or in a condition where no person received the outcome. (**B)** Choice behaviour in the three learning conditions, self (blue, prosocial (green) no one (orange). (**C)** Comparison of learning rates (LR) from the computational model. Participants had a significantly higher learning rate when learning for self compared to the prosocial and no one condition. (**D)** Individual differences in empathy modulated the prosocial vs. self learning rate difference, with those higher in empathy having a more similar learning rate between the prosocial and self conditions.

6.2.4. *Functional imaging and analysis*

6.2.4.1. fMRI data acquisition

A Siemens Avanto 1.5-T MRI scanner was used to acquire a 5.5-minute 3-dimensional T1-weighted structural scan and 424 multislice T2*-weighted echo planar volumes with blood oxygenation-level–dependent (BOLD) contrast. The structural scan was acquired using a magnetization prepared rapid gradient echo (MPRAGE) sequence. Imaging parameters were: 176 slices; slice thickness=1 mm; gap between slices=0.5 mm; TR=2730 ms; TE=3.57 ms; field of view=256 mm x 256mm2; matrix size=256 x 256; voxel size=1×1×1 mm resolution. The functional imaging sequence was acquired in an ascending manner, at an oblique angle (≈30˚) to the AC-PC line to decrease the impact of susceptibility artefact in the orbitofrontal cortex (Deichmann et al., 2003) and had the following acquisition parameters: 424 volumes, 1 mm gap; echo time=50 ms; repetition time=2975 ms; flip angle=90°; field of view=192 mm; matrix size=64x64.

6.2.4.2. fMRI data analysis

Imaging data were analysed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Data preprocessing followed a standard sequence: the first 4 volumes and last volume were discarded. Images were then realigned and co-registered to the participants own anatomical image. The structural image was processed using a unified segmentation procedure combining segmentation, bias correction, and spatial normalization to the MNI template using the New Segment procedure (Ashburner & Friston., 2005); the same normalization parameters were then used to normalize the EPI images. Voxel size was resampled to 1.5 x 1.5 x 1.5mm. Lastly, a Gaussian kernel of 8 mm FWHM was applied to spatially

smooth the images. Before the study, first-level design matrices were checked to ensure that estimable GLMs could be performed with independence (or 'rank sufficency') between the parametric regressors (Chosen value and PE in the three conditions), with correlations coefficients of r <0.25.

6.2.4.3. First-level analysis

Seven event-types were used to construct regressors in which event timings were convolved with the canonical haemodynamic response function. The three conditions at the time of the cues and 3 conditions at the time of the outcome were modelled as separate regressors. Each of these regressors was associated with a parametric modulator taken from the computational model. At the time of the cue this was the chosen value, and the PE at the time of the outcome. The instruction cue at the beginning of each block was also modelled. An additional regressor modelled missed trials where participants did not select one of the two symbols in the response window. For those participants where there was visible head motion in a particular scan (>1mm or 1 degree between one volume and the next) an extra regressor was included. These images were removed and replaced with an image created by interpolating the two adjacent images in order to prevent distortion of the between-subjects mask (4 participants, less than 1% of total time series). Six head motion parameters modelled the residual effects of head motion as covariates of no interest.

6.2.4.4. Second-level analysis

Contrast images from the first level were input into a second-level flexible-factorial design with one factor (PE) and three levels (self PE, prosocial PE, no one PE). Main effects are reported at p< .05, family-wise error (FWE) corrected at the cluster level across the whole brain or p<.05 small volume corrected in a priori regions of defined using structural masks taken from the appropriate anatomical atlas (ventral striatum, subgenual ACC, ACCg, bilateral DLPFC, and bilateral lateral OFC; toolboxes: Harvard-Oxford Atlas, regions 46v and 9 from

(Sallet et al., 2013) Anatomy toolbox regions s24 and 25, region 24 from (Beckmann, Johansen-Berg, & Rushworth, 2009) orbitofrontal cortex from AAL atlas, respectively). We also conducted a conjunction analysis (self PE ^ prosocial PE ^ no one PE) to identify regions that responded in all three conditions.

6.3. Results

6.3.1. *Behavioural results*

Analysis of total amount of money won between conditions showed that there was a significant main effect condition (F (2,60) = 3.2, p < .05). Post-hoc analyses showed a significant difference between money won in the self and no one condition only (M = 3165 points vs. 2994 points, p < .05). We also examined participants' choice behaviour, which showed a main effect condition (F (2,60) = 5.4, p < .01). Post-hoc analyses showed that participants selected the high probability option significantly more often for self compared to no one (M = .84%, SD= .02 vs. M = .77%, SD = .03, p= .02) and significantly more often for prosocial compared to no one conditions (M = .83%, .02 vs. .77%, SD= .03, p = .03) but there was no difference between self and prosocial conditions (p > .05). (**Figure 1, panel B**).

Analysis of the difference in LR parameters between the conditions showed a significant main effect of condition (F(2,60)=11.47, p< .001). Post-hoc analyses showed a significant difference in learning rate between self and prosocial (M = 3.7, p < .001) and self and no one (M = 3.2, p= .02) but no difference between prosocial and no one (M = 3.4, p = .53) (**Figure 1, Panel C**). These data suggest that people have a self bias in their learning rates, such that they learn faster about outcomes for themselves compared to when learning in the prosocial condition or no one. Analysis of the difference in temperature parameters between the three conditions also showed a significant main effect

(*p*<.01, Huynh-Feldt correction for sphericity violations). Post-hoc analyses showed a significant difference in the temperature parameter between the no one and self condition (M = .53 vs. .25, *p* = .05) and the no one and prosocial condition (M = .53 vs. .20, p<.01) but no significant difference between the self and prosocial condition (M = .25 vs. .20, *p* >.05) .This means that choices were made equally consistently for the self and prosocial conditions, but that people were more variable in their choice behaviour when selecting choices where no one received the outcome.

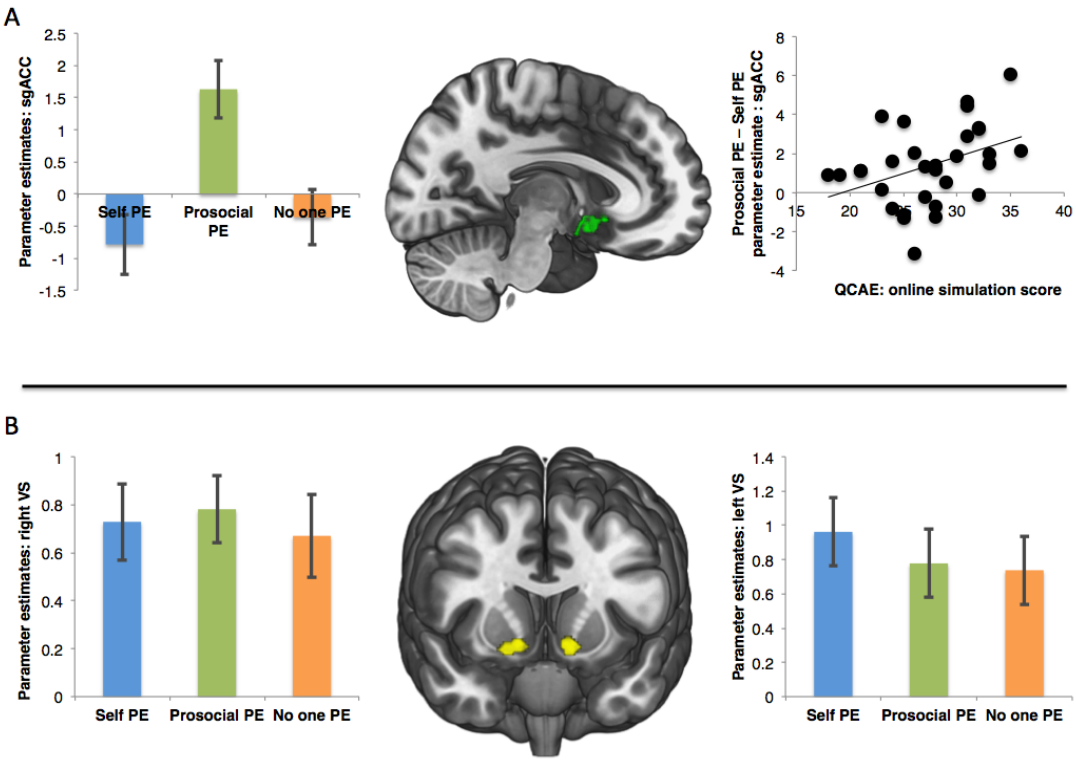
The second aim was to examine whether differences in learning rates and beta parameters between self and other were associated with individual differences in empathy. I observed that the online-simulation subcomponent of the QCAE was significantly positively associated with the prosocial-self learning rate difference (*r* = .41, *p* = .01). This suggests that individuals higher in online-simulation had a more similar learning rate between self and prosocial compared to those low in online-simulation (see **Figure 1 panel D**).

6.4. fMRI results

6.4.1.1. fMRI data: conjunction analysis to identify common coding of PEs

The ventral striatum responded to the three—way conjunction between self PE prosocial PE and no one PE (MNI coordinates [x=10, y=15, z=-9] Z=4.09, k=91, *p*< .01 SVC and [x=12, y=10, z=-11], Z=3.72, k=78 *p* = .023 SVC) (**See Figure 2, Panel B)**. Analysis in SPSS using the parameter estimates of the extracted clusters showed that responses in all three conditions were significantly above 0 (all ps<.01). No other regions responded significantly to the three-way conjunction whole-brain FWE cluster corrected level or in the ROIs.

Figure 6.2. Neural response in the sgACC and ventral striatum to PEs



Notes: (**A**) the sgACC encoded prosocial PEs exclusively and this response was modulated by individual differences in empathy. (**B**) The ventral striatum responses to PEs regardless of the agent the outcome was to be received by. Images displayed at p<.001 uncorrected.

6.4.1.2. fMRI data: contrasts between conditions to identify distinct coding of PEs

6.4.1.3. Prosocial PE > Self PE and No one PE

The sgACC showed a significant response to prosocial PEs exclusively (MNI coordinates [x=-2, y=4, z=-15] Z=3.83, k=148, *p*=.02 SVC) (see **Figure 2, Panel A).** Analysis in SPSS using the parameter estimates of the extracted clusters

showed that only the response in the prosocial condition showing a significantly different response from 0 (t(30)=4.3, p<.001). No other regions responded significantly in our ROIs or FWE whole brain corrected.

6.4.1.4. Self PE + No one PE > Prosocial PE

I observed opposing coding of self and prosocial PEs in left DLPFC (MNI coordinates [x=-36, y=18, z=43] Z=4.47, k=62, *p* < .01 SVC) and opposing coding of no one and prosocial PEs in right DLPFC (MNI coordinates [x=32, y=15, z=39] Z=4.36, k=27, *p* < .02 SVC). Analysis in SPSS using the parameter estimates of the extracted clusters showed that the response in the left DLPFC were significantly different from 0 for both the prosocial (*p* < .01) and self conditions (*p* < .001) but not in the no one condition (*p* = .16) (**See Figure 3, panel A**). We also observed opposing coding of self and prosocial PEs in the right OFC (MNI coordinates [x=34, y=48, z=-11] Z=3.82, k=81, *p* < .03 SVC) and opposing coding of no one and prosocial PEs in the left OFC (MNI coordinates [x=-34, y=54, z=14] Z=3.47, k=36, *p* = .08[marginal] SVC) (See **Figure 3, panel B**). One sample t-tests showed that the response to prosocial (*p* = .04) and self PEs (*p* = .04) in right OFC was significantly different 0 but not for no one PEs (*p* = .20). In the left OFC only the response to prosocial PEs was significantly different from 0 (*p* < .01). No other regions responded significantly in our ROIs or FWE whole brain corrected.

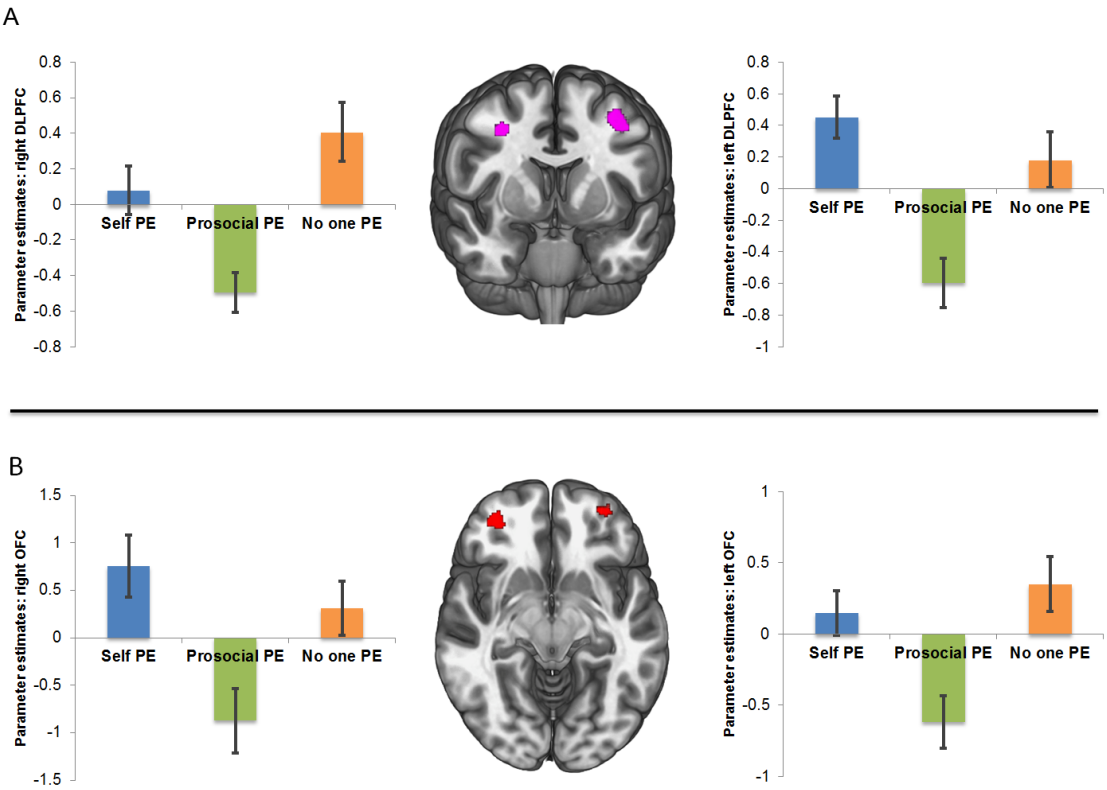
6.4.1.5. fMRI data: associations with trait empathy

To test our second hypothesis, that the extent to which these regions would signal prosocial PEs would be positively associated with trait empathy, I used MarsBaR (Brett et al., 2002) to extract each cluster contrast estimates from the neural regions defined above, and correlated these with participants' self-reported empathy on the five QCAE subscales.

Online simulation, the same empathy subscale that modulated differences in learning rates between self and prosocial conditions also modulated the

prosocial compared to self PE signalling in the sgACC (r = .39, p < .05). Emotion contagion modulated prosocial compared to self PE signalling in the right DLPFC (r = -.36, p < .05). No other subscales of the QCAE significantly correlated with the prosocial-self PE parameters (all ps > .06).

Figure 6.3. Neural response in the DLPFC and OFC to PEs



Notes: (A) The DLPFC showed opposing coding prosocial PEs relative to no one PEs (right DLPFC) or self PEs (left DLPFC). **(B)** The OFC also showed opposing coding of prosocial relative to PEs relative to no one PEs (right DLPFC) or self PEs (left DLPFC). Images displayed at p<.001 uncorrected.

6.5. Discussion

Reinforcement learning (RL) theory provides a powerful framework to understand how individuals learn to be prosocial. This framework can be used to understand the links between prosocial learning and neural response across species and in different learning contexts. In the current chapter I used RL

theory as a model to identify neural regions that signal prosocial prediction errors and whether these overlap with regions that signal PEs for ourselves. Second, I examined how individual differences in empathy modulate prosocial learning and neural response. I found that the subgenual cingulate cortex (sgACC) encoded prosocial PEs exclusively, and was the only region to do so. The ventral striatum responded to PEs in all three conditions, that is, regardless of the agent that received the outcome. Both left DLPFC and right OFC showed opposing coding of PEs for self and other, with these regions positively scaling with self PEs and negatively scaling with prosocial PEs. Individual differences in empathy modulated both behavioural and neural responses, with those higher in empathy having more similar learning rates between self and other, increased responsivity of the sgACC to prosocial relative to self PEs and more similar response of the right DLPFC to prosocial relative to self PEs. Together, these findings suggest that there are common and distinct neural regions that process rewarding outcomes for ourselves and other people and that individual differences in empathy can modulate the rate at which we learn about rewards for others as well as neural encoding of prosocial PEs. This provides a potential mechanism linking individual differences in empathy to variability in prosocial learning behaviour.

Previous studies have suggested a specific role for the sgACC in decisions to donate money to charity compared to decisions about monetary reward for ourselves (Moll et al., 2006). This region has also been consistently observed in paradigms of allocating rewards to others (Behrens et al., 2008; Hsu, Anen, & Quartz, 2008), in moral scenarios (Wiech et al., 2013; Zahn, Moll, et al., 2009), to vicarious reward (Mobbs et al., 2009) and shows increased responsivity in individuals higher in empathic concern (Zahn, de Oliveira-Souza, Bramati, Garrido, & Moll, 2009). The sgACC is connected to other regions involved in reward processing and social decision-making, including the nucleus accumbens, amygdala and orbitofrontal cortex (Johansen-Berg et al., 2008; Rushworth et al., 2007). We provide the first evidence that this region processes a prediction error that scales with the unexpectedness of outcomes - exclusively for another person.

As compared to other portions of the cingulate cortex, comparatively little is known about the role of the sgACC for social behaviour. This is in part because it is very difficult to record from this region in non-human primates or to cause focal lesions, as lesions to the vmPFC also often cause damage to the subgenual ACC, adjacent portions of orbitofrontal cortex and the dorsal ACC (Hadland, Rushworth, Gaffan, & Passingham, 2003). In one of the few studies to examine the effect of sgACC lesions, Rudebeck and colleagues tested the contribution of sgACC to sustaining autonomic arousal associated with positive emotional events (Rudebeck et al., 2014). The authors used a Pavlovian conditioning procedure where it had been shown that autonomic arousal increases in response to cues that predict rewards, and this arousal is maintained during an interval before a reward is delivered. They showed that although monkeys with lesions of the sgACC showed an initial, cue-evoked arousal, they failed to sustain this arousal until the anticipated reward was delivered. Rudebeck and colleagues thus suggest that the sgACC may contribute to positive affect by sustaining arousal in anticipation of positive emotional events. Other studies have implicated the sgACC and adjacent septal area in affiliative behaviours in human and animal studies (Depue & Morrone-Strupinsky, 2005; Insel & Young, 2001). These data suggest that this region encodes prosocial PEs, increasing its response when outcomes for others are unexpectedly positive and decreasing response when outcomes are unexpectedly negative.

In contrast, the ventral striatum responded to PEs in all three conditions, regardless of who the reward was to be received by. This supports existing studies showing ventral striatum response to self and other reward and PEs (Harbaugh et al., 2007; Moll et al., 2006; Morelli et al., 2015; Sul et al., 2015) but extends these findings to show that ventral striatum signals are also apparent when rewards are not directed to a specific agent. Such a profile is suggestive of this region coding a general learning mechanism regardless of the agent a reward is to be received by.

Both the left DLPFC and right OFC showed a pattern of opposing coding of self and prosocial PEs, i.e. increasing response to unexpectedly positive outcomes

for self, but decreasing response to unexpectedly positive outcomes for the other person. The DLPFC has previously been shown to respond in paradigms where other-regarding and self-regarding preferences come into conflict (e.g. (Sanfey, 2007), and thus has been suggested to be of importance in balancing prosocial and selfish behaviour (Hunt & Behrens, 2011; Sanfey, 2007). The observed opposing coding scheme for self and prosocial PEs is consistent with these findings and could have consequences for social behaviour, such as increased competitiveness in social interactions. I also found that those individuals who reported higher levels of empathy showed a more similar response of the right DLPFC to prosocial and self PEs. This suggests that the negative coding of prosocial PEs in right DLPFC is reduced in individuals higher in empathy.

The OFC is argued to be important for reinforcers to influence behaviour (Rushworth et al., 2007) and has been implicated in emotion, reinforcement learning and social behaviour. Individuals with lesions to the OFC exhibit changes in social behaviour (Willis, Palermo, Burke, McGrillen, & Miller, 2010) and neurons in the OFC may encode the motivational value of social information (Watson & Platt, 2012). The current results suggest that this region may distinguish self and other outcomes by encoding PEs for oneself and for another person in opposite directions.

Behaviourally I observed that participants showed a self bias in their learning rate, namely that they had a higher learning rate for themselves compared to the prosocial or the no one condition. Intriguingly, individual differences in online simulation, a key component of empathy, modulated this response with those higher in online simulation showing a more similar learning rate between self and other than those lower in online simulation. The finding that individual differences in empathy can modulate the rate at which we learn about rewarding outcomes for others suggests a potential mechanism by which empathy could facilitate prosocial decisions, namely by increasing the sensitivity to rewarding outcomes for other people. I also found that it was this same component of empathy, online-simulation - the ability to put oneself in another person’s position by imagining what that person is feeling (Reniers et al., 2011)

- that modulated response of the sgACC to prosocial PEs. Previous studies have suggested that empathy can modulate how we process rewards for other people in a temporal discounting task (O'Connell et al., 2013). I suggest that conceptualising empathy within a framework of reward sensitivity could be fruitful in further studies examining how individual differences in social cognition translate into individual differences in choice behaviour.

I did not observe response of the ACCg in any of the conditions. This may seem surprising given that this region has previously been implicated in processing rewards for others and in social behaviour (Apps, Green, et al., 2013; Apps, Lockwood, et al., 2013; Apps & Ramnani, 2014; Behrens et al., 2008; Boorman et al., 2013; Chang et al., 2013; Rudebeck et al., 2006). Whilst it is notoriously difficult to interpret a null finding as there can be a number of factors for not observing a predicted neural response, one possible explanation is the role of 'reference frames' in social decision-making (e.g. Hunt & Behrens, 2011). In this task participants believed that their choices for the other participant were anonymous and they were not being observed when making decisions. Thus, these decisions were not made in the same reference frame as previous studies of ACCg that have examined social interaction or observing another's responses. A key test of this hypothesis would be to compare a condition where the selecting of rewards for others was anonymous compared to known to see whether when the set up of this paradigm is posed as a social interaction ACCg response is observed.

Successful social behaviour and empathic abilities are thought to be compromised in a number of disorders including psychopathy and autism (Bird & Viding, 2014). The paradigm developed here could be used in further studies examining how individuals with disorders of social behaviour learn about rewards for other people and neural regions that signal self and prosocial PEs. This could help to identify mechanisms that link disturbances in empathy to individual differences in social behaviour. Conceptualising prosocial learning within a RL framework also allows for the current paradigm to be extended to investigate prosocial learning across species.

6.5.1. *Conclusions*

In summary, I identified key neural regions that signal prosocial prediction errors and how these overlap and are distinct from regions that signal prediction errors for ourselves. I also show that individual differences in the ability to take the perspective of another person, a component of empathy, can modulate both the rate at which individuals learn about rewards for others and neural responses in the sgACC. Taken together, these findings provide evidence to support both the common currency and socially specific accounts of social decision-making and provide a potential mechanism linking empathy to prosocial behaviour.

In the final chapter, I summarise the findings of the preceding empirical chapters and discuss the implications and future directions of this work.

CHAPTER 7: General discussion

7.1. Overview

Empathy, the ability to vicariously experience and to understand the affect of other people (Bird & Viding, 2014; Decety & Jackson, 2004; Eisenberg, 2000; Hoffman, 2001; Singer & Lamm, 2009), is a key ability for successful social cognition and behaviour. Consequently, it is not only crucial to define the component processes of empathy carefully, but also to systematically identify the behavioural and neural mechanisms by which vicarious experience influences social cognition and behaviour.

At present there is still much to find out about different processes involved in empathy. It is often suggested that empathy is a multidimensional phenomenon (Bird & Viding, 2014; Decety & Jackson, 2004; Eisenberg, 2000, Hoffman, 2001; Singer & Lamm, 2009). However, few studies have examined different components of empathy within the same sample. Moreover, when investigating individual differences in disorders associated with reduced empathy, such as psychopathy and autism, many researchers have employed tasks that draw on multiple processes implicated in empathy. This can make it difficult to isolate specific difficulties related to one disorder or another. By using paradigms that separate different processes clearly, we can perhaps more precisely delineate the processing atypicalities associated with specific disorders.

Vicarious experience has often been suggested to facilitate positive social behaviours, yet existing paradigms have often focused on examining associations between concern (e.g. sympathy) for others and prosocial tendencies (e.g. Batson, 1998). There is a lack of empirical data that has identified whether the vicarious perception of another's experience is a motivating factor for prosocial behaviour. It has been proposed that antisocial behaviours in children with conduct problems could stem from disrupted

empathic processing (Blair, 2005), yet we are still to identify whether individuals with antisocial behaviour display atypical neural responses to the suffering of others. Finding whether individuals with conduct problems have a reduced neural response to other people's distress could help us understand one potential mechanism linking empathy to prosocial behaviour.

Effectively cooperating or competing with another requires the ability to compute the value of stimuli that predict positive experiences, such as rewards, for others (Ruff & Fehr, 2014). However, very little is known about how vicarious reward predictions are processed in the brain or how individual differences in social functioning are related to neural response to others' reward. Many of the studies to date have instead focused on the perception of others' negative experiences, but it is important to understand whether similar neural responses characterise positive experiences.

Finally, research has begun to apply well-characterised computational models of reward learning to reward learning in social contexts (Behrens et al., 2008; Burke et al., 2010; Hampton, Bossaerts, & Doherty, 2008; Ruff & Fehr, 2014). Using these mathematical models can help us to examine the specific neural computations that link empathy to prosocial behaviour and neural responses in different learning contexts and across different species.

The current thesis set out to advance knowledge in these key areas using a multimodal approach of questionnaires, behavioural paradigms, neuroimaging and computational modelling, in both typical samples with variability in empathy and in children with conduct problems and low levels of empathy.

7.2. Research questions

To summarise, four outstanding research questions regarding relationships between empathy, social cognition and behaviour were identified:

1) **How separable are different aspects of empathic processing and do distinct aspects of empathic processing related to different aspects of social functioning?**

2) **How does empathy relate to trait prosocial behaviour and do additional trait constructs moderate the relationship between empathy and prosocial behaviour? (Chapter 3)**

3) **Where in the brain is vicarious information (both negative and positive) processed, and does this vary in: (i) children with conduct problems and with (ii) individual differences in typical empathy? (Chapters 4 and 5)**

4) **Which neural regions signal prosocial prediction errors? What are the mechanisms that link empathy to prosocial decision-making behaviour and neural response? (Chapter 6)**

Findings and implications pertaining to each of these questions are considered sequentially in the sections below. Overall, it is argued that empathy is one of the key processes that can aid in successful social functioning and explains important variability in social cognition and behaviour.

7.2.1. *How separable are different aspects of empathic processing and do distinct aspects of empathic processing related to different aspects of social functioning?*

In **Chapter 2** I presented a study that used two behavioural paradigms measuring key components of empathy - affective resonance and cognitive-perspective taking - and collected trait measures of psychopathic, autistic and alexithymic traits in the general population. I found that performance on these two tasks was significantly positively associated, such that those who scored higher in affective resonance also scored higher in cognitive-perspective taking and vice versa. I also found that individuals with high levels of psychopathic and autistic traits were characterised by difficulties in different components of empathy. Individuals with high levels of psychopathic traits showed reduced resonance with others' emotions, but not with cognitive perspective taking. Conversely, individuals with high levels of ASD traits showed problems with cognitive perspective-taking but not with resonating with others' emotions. This suggests that although behaviourally individuals with psychopathy or ASD may appear to lack empathy, this could be for different reasons. A second aim of this study was to examine whether trait levels of alexithymia were able to explain empathy impairments associated with either psychopathy or ASD dimensions. Alexithymia was found to be associated with problems with affective resonance, but this association was independent of psychopathic traits, suggesting that different component processes (reduced tendency to feel what others feel and reduced ability to identify and describe feelings) could comprise performance on the affective resonance task. Alexithymia was not associated with the reduced cognitive perspective-taking and could not thus account for the association between reduced cognitive perspective-taking and higher levels of ASD traits.

Overall, the findings of this study suggested that although affective resonance and cognitive perspective-taking measures share some variance (future studies could explore candidate processes that may influence performance on both measures, such as executive functioning), they can capture dissociable processes, and thus extends our knowledge regarding the structure of empathy.

Moreover, elevated psychopathic and ASD traits are characterised by difficulties in *different* social information processing domains. This finding supports and extends previous work in children with conduct problems and callous-unemotional traits compared to children with ASD that has also showed different profiles of empathy impairments (Jones et al., 2011; Schwenck et al., 2012).

This work also extends our knowledge of the association of alexithymia with psychopathic and autistic traits (e.g. Lander et al., 2012; Louth et al., 1998). The alexithymia hypothesis states that, where observed, reduced empathy in individuals with ASD may be attributable to co-occurring alexithymia rather than autism per se (e.g. Bird & Cook, 2013). I show that reduced affective resonance in individuals with elevated psychopathic traits and reduced cognitive perspective taking in individuals with elevated ASD traits are not explained by co-occurring alexithymia. However, alexithymia is independently associated with reduced affective resonance over and above psychopathic and ASD traits. This suggests that reduced cognitive empathic processing in ASD does not appear attributable to alexithymia and that to the extent that individuals with ASD show difficulties in affective resonance, this is likely to be due to co-occurring alexithymia (as outlined in Bird & Cook, 2013), but note that I did not find high levels of ASD traits in the typical population to be associated with difficulties in affective resonance.

Limitations and future directions

These findings will need to be assessed in a clinical sample to examine whether similar associations with alexithymia occur. Moreover, further studies could examine a wider battery of associated processes related to empathy such as mimicry, empathic concern and identification of emotions. It will also be of interest to determine whether the processing atypicalities associated with psychopathic, ASD and alexithymia traits explain real life observations of unempathic behavior, as rated by others or observed in an experimental setting.

7.2.2. *How does empathy relate to trait prosocial behaviour and do additional trait constructs moderate the association between empathy and prosocial behaviour?*

In **Chapter 3** I presented a study that examined associations between components of empathy and prosocial tendencies, and whether individual differences in the ability to regulate one's own emotions moderated associations between empathy and prosocial behaviour. I found that affective and cognitive empathy predicted self-reported prosocial tendencies. Specifically, these components were able to explain both joint and unique variance in prosocial tendencies. In addition, cognitive reappraisal moderated the association between affective empathy and prosocial tendencies. Whilst there was a significant positive association between empathy and prosocial tendencies for individuals with a low or average tendency to reappraise there was not a significant association for those with a high tendency to reappraise.

These findings suggest that, in general, empathy is positively associated with prosocial behaviour and that both affective and cognitive dimensions of empathy could be important for either motivating prosocial tendencies or prosocial behaviours that are seen as socially appropriate in certain situations. However, for individuals who have a high tendency for cognitive reappraisal the association between empathy and prosocial behaviour may not be as strong. I suggest that this could be because those individuals with a high tendency to reappraise a situation may be more readily able to deduce the desirability of prosocial behaviours without a vicarious emotional experience. This work extends that of existing studies that have linked empathic concern to prosocial motivation (e.g. Batson, 1998; Davis, 1983) and highlights that vicarious experience can also facilitate prosocial tendencies.

Limitations and future directions

One limitation of this study is the reliance on self-report. Future studies with behavioural paradigms, such as the paradigm developed in **Chapter 6** or

economic games such as the dictator game or ultimatum game, will need to be used to examine whether such self-reported tendencies translate into differences in behaviour.

7.2.3. *Where in the brain is vicarious information processed, and does this vary in:*

7.2.3.1. Children with conduct problems

In **Chapter 4** I examined how individual differences in empathy impact upon negative social behaviour, that is, antisocial behaviour. I used functional magnetic resonance imaging (fMRI) to examine neural responses to others' pain in a sample of children with conduct problems, who show high levels of antisocial behaviour, and varying levels of callous and unemotional traits. Neural responses were compared to a control group matched for IQ, age, socioeconomic status and ethnicity. I also collected parent and teacher ratings of conduct problem symptoms and callous and unemotional traits to examine individual differences in neural response. I found that, relative to controls, children with conduct problems showed reduced blood oxygen level-dependent responses to others' pain in bilateral anterior insula (AI), anterior cingulate cortex (ACC), and inferior frontal gyrus, regions associated with the vicarious perception of others pain in previous studies (Fan et al., 2011; Lamm et al., 2011). I also observed that in the conduct problem group, callous traits were negatively associated with responses to others' pain in AI and ACC, such that those highest in callous traits showed the greatest reduction in neural response of these regions. Additional analyses showed that the gyral portion of the ACC, ACCg, was one of the subregions of the ACC where these differences between children with CP and typical children were observed. These findings suggest that children with CP have atypical neural responses to others' pain. The negative association between callous traits and AI/ACC response could reflect an early neurobiological marker indexing risk for empathic deficits seen in adult psychopathy.

Limitations and future directions

An outstanding question is whether individuals with CP have an atypical experience of pain themselves. Given that I did not have a self pain condition I was unable to assess the profile of self pain processing in children with CP. Preliminary evidence suggests that children with CP may indeed have disrupted self pain processing (Cheng et al., 2012). Future studies could examine the processing of pain in children with conduct problems, if this could be done in an ethically feasible manner, to test whether the reduced neural response to others' pain is also reflected by a reduced general response to pain. One option would be to apply painful stimuli that were painful enough to be aversive but not to cause harm.

7.2.3.2. Individual differences in typical empathy

In **Chapter 5** I described a study**,** which focuses on the vicarious perception of positive experience, in particular other peoples' reward. I used fMRI to examine neural responses to cues that signalled the likelihood of other's reward in a key neural region thought to encode vicarious experience, the ACCg. Importantly, this paradigm used a social confederate rather than pictures of other people, which may help to uncover neural mechanisms that support social interactions. I also measured individual differences in empathy to examine how these differences modulated neural response. I found that the ACCg robustly signalled the likelihood of a reward being delivered to another. In addition, this ACCg response significantly co-varied with trait emotion contagion, a necessary foundation for empathising with other individuals. In individuals high in emotional contagion the ACCg was specialised for processing others' rewards exclusively, but for those low in emotion contagion this region also responded to information about the subject's own rewards.

These results are the first to show that the ACCg signals probabilistic predictions about rewards for other people, and that the substantial individual

variability in the degree to which the ACCg is specialised for processing others' rewards is related to trait empathy. The response of ACCg to others' rewards fits with a large body of evidence showing involvement of the ACCg in social cognition and behaviour across species (Rudebeck et al., 2006; Behrens et al., 2008; Apps, Lockwood et al., 2013, 2015; Boorman et al., 2013; Chang et al., 2013; Apps & Ramnani, 2014; Jones et al., 2011). The association between ACCg response and emotion contagion suggests that variability in empathy may not only be associated with the extent to which social information is processed in ACCg, as suggested in previous studies and theoretical accounts of the neural mechanisms of empathy (e.g. Lamm et al., 2011; Engen and Singer, 2013), but also the extent to which self as well as other reward information is computed. These findings may help reconcile previous discrepancies in the functions that have been imputed to the dACC in terms of competitive social behaviours (Hillman & Bilkey, 2012; Haroush & Williams, 2015) but also empathy (Lamm et al., 2011; Engen and Singer, 2013).

Limitations and future directions

One limitation of this study was that the behavioural component of the task was limited. Although participants were required to judge whether the outcomes for themselves and the other person were expected or unexpected, and this effect was associated with emotion contagion (those highest in emotion contagion showed the greatest speeding of response when judging outcome for the other person) this behavioural component was not at the point of interest for neural response (the judgment at the outcome rather than at the cue). Consequently I cannot assess how ACCg response in this study translated into behaviour.

7.2.3.3. Which regions of the brain are involved in signaling prosocial prediction errors? Do individual differences in empathy predict variability in prosocial learning and its neural basis?

In **Chapter 6** I present a study that adopted a multimodal approach of combining questionnaire measures of trait empathy, computational modelling of behaviour and brain imaging to uncover neural regions that signal prosocial prediction errors and variation with trait empathy. Behaviourally, I found that participants showed a self bias, that is they had a higher learning rate when learning about rewards delivered to themselves compared to other people or neither person. Trait empathy modulated the prosocial-self learning rate difference such that those higher in empathy had a more similar learning rate between self and other. Neurally, I found that key neural regions involved in decision-making about other's rewards, namely ventral striatum, sgACC, OFC and DLPFC signalled prosocial PEs. The ventral striatum signalled PEs regardless of the agent they were to be received by. The sgACC was the only region to code prosocial PEs exclusively and the left DLPFC and right OFC showed opposing coding of self and prosocial PEs, increasing response to unexpectedly positive outcome for self and decreasing response to unexpectedly positive outcomes for the other participant.

The observation that individual differences in empathy can modulate the rate at which we learn about rewarding outcomes for others suggests a potential mechanism by which empathy could facilitate prosocial decisions, namely by increasing the sensitivity to rewarding outcomes for other people. The response of the ventral striatum in all three learning conditions suggests that this region may code a general learning mechanism, regardless of agent, rather than response only to self and other reward. This supports previous studies showing ventral striatum response to both self and social prediction errors (Harbaugh et al., 2007; Morelli et al., 2015; Sul et al., 2015; Moll et al., 2006) but shows that the ventral striatum signal may related to a general learning mechanism rather than a response that codes specifically for self and other rewards. The modulation of prosocial prediction errors in the sgACC exclusively points to an important role of this region in signalling unexpected outcomes for others following our actions, suggesting a potential mechanism for why sgACC response has been linked to prosocial decisions in previous studies (e.g. Behrens et al., 2008; Wiech et al 2013; Zahn et al., 2009) and in vicarious reward (Mobbs et al., 2009). This region also correlated with online simulation

(a component of empathy), with those highest in online simulation showing the greatest response to prosocial PEs relative to self PEs. Taken together, these findings show that empathy can modulate both behavioural and neural responses to prosocial decisions, and identifies key neural regions that may be involved in signalling prosocial PEs. These findings also suggest that both 'socially-specific' and 'common currency' (c.f. Ruff and Fehr, 2014) neural regions are important for signalling social decisions.

7.3. Implications and future directions

The findings from this thesis have generally supported the claim that that empathy is important for successful social functioning and meaningfully relates to individual differences in social cognition and behaviour. **Chapters 2** and **3** showed that empathy was differentially related to trait variability in disorders of social functioning and prosocial tendencies, respectively. **Chapters 4, 5** and **6** showed that individual differences in empathy explained variance in neural responses during vicarious perception and decision-making. In the next section I discuss the implications and future directions of these findings.

7.3.1. *Implications for understanding individual differences in social behaviour*

7.3.1.1. Psychopathy and autism spectrum disorders (ASD)

Psychopathy and ASD have both been suggested to be characterised by reduced empathy, although it has been argued that this is likely for different reasons (Bird & Viding, 2014). **Chapter 2** adds growing support to the claim that traits associated with these disorders are indeed related to different aspects of empathy (e.g. Blair et al., 2005; Bird & Viding, 2014). I found that alexithymia explained reduced affective resonance over and above psychopathic traits,

suggesting that different component processes, the tendency to feel what others feel and the tendency to identify and describe feelings, comprise the construct of affective resonance. Moreover, these findings suggested that the affective impairments in individuals high in psychopathic traits are not related to alexithymia. Interestingly, alexithymia did not explain variance in reduced cognitive perspective-taking in individuals high ASD traits. Further studies in clinical samples with psychopathy will be needed to examine whether this same pattern occurs at clinical levels, and to further examine whether alexithymia can explain reduced cognitive perspective-taking in individuals with autism, given that alexithymia is argued to account for reduced affective resonance in these individuals (e.g. Bird & Cook, 2013).

However, given that both psychopathy and ASD are associated with atypical social cognition, there may be neural regions that are similarly disrupted in the two disorders. For example, the ACCg is a key candidate for common atypicalities in the two disorders. Previous studies have suggested that a similar portion of the dACC that was activated in **Chapters 4 and 5** is anatomically and functionally atypical in individuals with psychopathy and in individuals with autism (e.g. Simms et al., 2009; Brazil et al., 2011; Anderson & Kiehl, 2012; Delmonte et al., 2013; Zikopoulos & Barbas, 2013; **Chapter 4**). One potential computation that is associated with ACCg function, and could be atypical in the two disorders, is reward processing.

Social Motivation Theory (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012; Dawson et al., 2004; Dawson, Webb, & McPartland, 2005) proposes that individuals with ASD are unable to form stimulus-reward contingencies for social stimuli, resulting in reduced social attention and engagement. Chevallier et al. (2012) suggested that these computations are underlied by an orbitofrontal-striatal-amygdala circuit. However, other evidence suggests that the ACCg may play a key role in the social impairments seen in ASD (Delmonte et al., 2013). Previous studies have shown disturbed cytoarchitecture specifically in the ACCg in individuals with ASD (Simms et al., 2009). Similarly, Delmonte et al. (2013) showed hyperconnectivity between the caudate and ACCg in children with ASD, the strength of which was negatively correlated with

neural responses to social rewards (Delmonte et al., 2012). Balsters et al. (2015) recently found that individuals with ASD showed reduced responding in ACCg to social reward prediction errors (Balsters et al., 2015). In their task, based on Apps et al., (2013), participants with autism and matched healthy controls were asked to choose to open doors that were probabilistically associated with a rewarding outcome either for themselves, for another person or a computer. At the time of the outcome participants were asked to judge whether the reward was expected or not. The authors suggest that this type of set up can be seen as similar to a false belief task, and prediction errors can be conceptualised as false beliefs, i.e. a person's recognition of unexpected outcomes for another person. Balsters and colleagues found that, behaviourally, individuals with autism were less able to judge unexpected rewards for another person or for a computer, but were not impaired when judging these outcomes for themselves. Neurally, they found a reduced response in the ACCg when individuals with autism saw unexpected outcomes for another person or a computer but not for themselves (Balsters et al., 2015). These preliminary studies suggest that the ACCg could play a role in coding social rewards and that this process is atypical in individuals with autism.

Similarly, atypical social reward processing has also been suggested characterise individuals with psychopathy, who are suggested to be insensitive to rewards that others will receive, leading to increased competitive behaviours (Curry, Chesters, & Viding, 2011; Koenigs, Kruepke, & Newman, 2010; Mokros et al., 2008) and decreased prosocial behaviours (Foulkes et al., 2014; White, 2014). Individuals with psychopathy have been shown to display reduced error related negativity, measured using Electroencephalography, when observing other's outcomes during a social interaction (Brazil et al., 2011). This signal is putatively sourced in the ACC. Recent studies also indicate that grey matter volume and activity in the ACCg correlate with psychopathic and callous traits (De Brito et al., 2009; Anderson & Kiehl, 2012; Cope et al., 2012; **Chapter 4**).

Taken together these studies suggest that the ACCg may be atypical in both psychopathy and autism. However, as shown in **Chapter 2** and many other studies these disorders are thought to be associated with *different* empathy

impairments (Bird & Viding, 2014; Blair, 2005; Jones et al., 2010; Schwenck et al., 2012). Consequently, whilst ACCg response may be atypical in the two disorders this could happen for different reasons. One potential explanation is the differences in the structure, function and connectivity of ACCg constrain the extent to which this region processes rewards for others compared to self, which may lead to atypical empathic processing, as suggested in **Chapter 4**. In **Chapter 4** I found that across participants the ACCg signalled reward prediction for other people but that in individuals low in emotional contagion this region also signalled information about rewards for self. The study by Chang and colleagues that recorded from neurons in the ACCg in non-human primates found that this region contained a higher proportion of 'other' selective neurons compared to 'self' selective and 'both' selective neurons (Chang et al., 2013). It is plausible that individuals with psychopathy and autism may have a different proportion of self, other and both reward responsive neurons in ACCg both compared to typical individuals, but compared to one another. This hypothesis would not be viable to test in a human population but it is conceivable that animal models could be used.

Another related possibility is that the connections between ACCg and other regions relate to distinct behavioural impairments seen in the two disorders. Further studies with individuals with psychopathy and ASD could use the paradigm developed in **Chapter 5** to examine vicarious reward perception, which was suggested to activate ACCg, and examine the connections between this region and others thought to be involved in social cognition and reward processing, such as dorsal regions of the medial prefrontal cortex, temporal poles, temporo-parietal junction and striatum (Barbas et al., 1999; Haber et al., 1995; Lynd-Balta & Haber, 1994; Markowitsch et al., 1985; Seltzer & Pandya, 1989; Yeterian & Pandya, 1991). This would allow the claim that these disorders are associated with atypical ACCg response to social information, but also more importantly to determine how an atypical response of ACCg could be associated with psychopathy and autism when these disorders are suggested to be characterised by separable empathy impairments.

It would also be of interest to test individuals with social disorders on these paradigms to help to resolve how empathy relates to reward sensitivity for others and prosocial learning, using the paradigm in **Chapter 6**. In **Chapter 6** trait empathy modulated the amount that people learnt about rewarding outcomes for themselves and another person, with this parameter being more similar in those higher in empathy. An insensitivity to other peoples rewards may lead to reduced prosocial behaviour, which is thought to characterise individuals with psychopathy (White et al., 2014; Foulkes et al., 2014). A recent study also showed that autistic traits modulated the association between social rewards and prosocial behaviour, but did not affect reward learning for oneself (Panasiti, Puzzo, & Chakrabarti, 2015). Together these findings suggest that both autism and psychopathy are related to insensitivity to rewards for other people. Consequently, determining the extent to which these individuals also have insensitivity to rewards themselves is an important avenue for future research. Measures of alexithymia could help to determine whether alexithymia can explain variance in reward learning for other people in both autism and psychopathy or whether comorbid alexithymia in autism is accounting for a reduced sensitivity to others' rewards. There is a preliminary suggestion that some associations between autistic traits and atypical social reward processing may indeed be accounted for by alexithymia (e.g. Foulkes, Bird, Gökçen, McCrory, & Viding, 2015)

Intriguingly, recent evidence from neuroimaging suggests that reduced neural responses to others' pain in individuals with psychopathy can be changed dependent on the instructions given to participants, and in particular if participants are explicitly instructed to empathise (Meffert et al., 2013). However, it remains to be seen whether effortfully activating the neural response to others' pain can foster empathy and empathic behaviour in individuals with high psychopathic traits. Nevertheless, factors that motivate effortful empathy in these individuals could be a key target for future research. Perhaps an explicit instruction to empathise during prosocial learning could facilitate performance in individuals with psychopathy and allow action-reward associations for other people to be formed. Another option could be to pair

outcomes for other people with outcomes for self to make outcomes for other people more motivating for individuals with psychopathy.

7.3.1.2. Healthy ageing

Another implication and avenue for further research is the extension of research investigating empathy and social behaviour to healthy ageing. There is suggestion that our level of social engagement can impact on our health and well-being across the lifespan, particularly in old age. By 2050 there will be an estimated 19 million people aged 65+ in the UK (Cracknell, 2010). In older adults, individuals who are less socially active are at higher risk of developing dementia (Fratiglioni, Paillard-Borg, & Winblad, 2004), general cognitive decline (James, Wilson, Barnes, & Bennett, 2011), memory loss (Ertel, Glymour, & Berkman, 2008) and motor function problems (Buchman et al., 2009). Therefore it appears that maintaining our social functioning is important for many aspects of healthy ageing.

However, there is also evidence that as we get older our social abilities can change (Moran, 2013). Cognitive aspects of empathy - as measured by tasks where participants infer others' mental states by viewing faces, cartoons, or stories - are typically associated with reduced task performance in older adults (reviewed in Moran, 2013). Importantly, such reductions are at least partly independent of any general cognitive decline (Kemp, Després, Sellal, & Dufour, 2012). In contrast, affective aspects of empathy and our willingness to engage in prosocial behaviours might even increase, or at least not be affected, by age (Beadle, Sheehan, Dahlben, & Gutchess, 2013). Very few studies have been conducted but the data suggest that older adults, compared to younger adults, report similar levels of trait affective empathy (Bailey, Ruffman, & Rendell, 2013; Beadle et al., 2013; Sullivan & Ruffman, 2004), divide money more generously in economic games (Bailey et al., 2013) and offer to donate more money to charities (Sze, Gyurak, Goodkind, & Levenson, 2012). The paradigms developed in this thesis could be used to examine the profile of social cognition

in healthy ageing to provide insight into strategies to scaffold healthy social ageing as well as the construct of empathy itself and associations with social behaviour.

7.3.1.3. Relationship of empathy to social behavior

Empathy is often assumed to be an important facilitator of prosocial behaviour, and the two processes are frequently linked. However, most of the current research investigating associations between these constructs has conceptualised prosocial behaviour as part of empathy (Zaki & Ochsner, 2012) or investigated associations between empathic concern and compassion or prosocial behaviour (Batson, 1998; Singer & Klimecki, 2014), rather than the association between vicarious experience and prosocial behaviour. This has made it hard to establish whether vicarious experience is indeed a motivating factor for individuals to behave prosocially. The findings of **Chapter 3** supported a link between vicarious experience and trait prosocial behaviour, with individuals reporting higher levels of both affective and cognitive empathy reporting higher prosocial tendencies. The findings of **Chapter 6** also supported a close association between the two constructs, with individuals higher in trait empathy showing a more similar learning rate between self and other.

7.3.2. *Implications for theory-theory and simulation theory*

In the introduction, I outlined that historically there were two prominent accounts of social cognition, theory-theory and simulation-theory. Theory-theory posited that we understand others' minds by forming a folk psychological theory, that is, a set of concepts about others (beliefs and desires) and governing principles as to how these concepts interact (e.g., people act to satisfy their desires according to their beliefs). Simulation-theory argued that we understand the minds of others via a process simulation, where the cognitive mechanisms that underpin the processing of one's own actions and intentions are simulated when processing the same information about another (Gallese and Goldman,

1998; Gallese, 2007). The discovery of 'mirror-neurons', neurons that increase their spike frequency during both action observation and action execution, has been invoked as evidence of 'mirror' like neural properties (Gallese and Goldman, 1998; Rizzolatti and Craighero, 2004; Gallese, 2007). In addition, the discovery of a core neural circuit that is engaged when processing others' mental states (Frith and Frith, 1999, 2006), is supportive of the theory-theory account. However, it is now generally accepted that these two accounts do not need to be mutually exclusive, and neither can sufficiently explain all aspects of social cognition on their own. Moreover, the account of mirror-neurons 'mirroring' has been convincingly challenged by accounts of associative-learning (e.g. Cook et al., 2014).

Nevertheless, it is important to understand the influence that these two accounts have had on the field of social cognitive neuroscience. In studies on vicarious perception and decision-making there are debates as to whether a neural region is processing social information if this same region also processes first-person information or whether a neural region is processing social information if it responds to third-person information exclusively (Engen & Singer, 2013; Ruff & Fehr, 2014). In this thesis evidence has been found supporting both types of neural responses in processing social information. In **Chapter 5** I showed that the ACCg, which has previously been shown to process social information, showed a response mainly to other reward and not self reward prediction. This fits with a large body of evidence suggesting relative specificity of ACCg for processing others' reward (reviewed in Apps, Lockwood, et al., 2013), which fits with theory-theory accounts of social cognition, i.e. that there are specific neural regions that process information about others. In **Chapter 6** neural regions were identified that both showed an overlap between social and non-social information (ventral striatum) as well as in regions that responded exclusively to social information (sgACC) or showed opposing coding between self and other (left dlPFC, right OFC). These findings support the claim that both socially specific and shared neural systems can process social information.

One advantage of having these theoretical frameworks, however, is in stimulating ideas for experimental paradigms and providing testable hypotheses. In an attempt to test the shared-representation hypotheses, i.e. a hypothesis inspired by the theoretical foundations of simulation theory, novel designs have been used borrowed from research in fields outside social neuroscience. For example, evidence from literature on pain has identified a placebo analgesia effect whereby individuals report pain reduction after being instructed they are being administered a potent painkiller, which is actually an inactive compound (e.g. Benedetti, Mayberg, Wager, Stohler, & Zubieta, 2005). Rutgen and colleagues recently applied this to examine overlap between self and other pain, the hypothesis being that if the shared representation account holds then the placebo analgesia effect should extend to other people's pain (Rütgen, Seidel, Riečanský, & Lamm, 2015). Rutgen and colleagues found that both self report measures of empathy and neural responses to others pain were affected by placebo analgesia (Rütgen et al., 2015), which they interpreted as supporting a shared functional overlap between self and simulating other pain. Other studies have tried to use advanced fMRI methods such as multivariate pattern analysis to identify if a region shows response in two separate domains, such as social and physical pain, whether the putative pattern of neuronal activity is in fact the same (e.g. Woo et al., 2014). Such a method can be applied to examine shared-representation interpretations of neural response where they are apparent. This method could also be applied to examine neural response to self and other pain and reward to test whether similar neural patterns response to both conditions and to examine whether overlapping neural response in the striatum to self, prosocial and no one PEs (identified in **Chapter 6**) show similar patterns of neuronal activity. We already know from the computational model that these regions are covarying with a PE signal, yet it is feasible that this could be differentially encoded in the different brain regions.

7.3.3. *Comparative studies and mathematical models*

The importance of comparative studies and borrowing well characterised theories and models of reward-guided behaviour, perception and decision-

making, which can be studied across species, will be important when moving forward with future research. The use of computational modelling techniques could help us link behavioural and brain processes that might be disrupted in psychopathy and autism but also help us understand how typical social decision-making operates. Youths with psychopathic traits have been found to show reinforcement reward learning impairments in the caudate and ventromedial prefrontal cortex (White et al., 2013). Specifically, White and colleagues found that children with conduct problems showed reduced response to positive prediction errors in the insula and decreased responses to negative PEs in the caudate. The study by White and colleagues was the first to use model-based fMRI in a sample of children with conduct problems. The advantage of this approach is that it can tell us not just where in the brain there might be differences between those with and without psychopathy but also how different cognitive process are implemented (O'Doherty, Hampton & Kim, 2007). The key question would be to examine how differences in reward learning for self and other translate into differences in empathic/prosocial behaviour by including a condition that examines reward outcomes for other people, such as in **Chapter 6**.

7.3.4. *Lesion studies*

In **Chapter 6** the sgACC was identified as the only region to signal prosocial PEs exclusively. Whilst many studies have investigated the function of the ACCg and ACCs, comparatively little is known about the function of the subgenual cingulate cortex region (sgACC, area 24b and area 25) for social behaviour. This is in part because it is very difficult to record from this region in non-human primates or to cause focal lesions, as lesions to the vmPFC also often cause damage to the subgenual ACC and adjacent portions of orbitofrontal cortex and the dACC (Hadland et al., 2003). However, the sgACC has been a target of deep brain stimulation for depression in patients (Mayberg et al., 2005) and it would be of interest to collect data on changes in social behaviour after stimulation of this region. Future studies could also examine how focal lesions in different neural regions, in both humans and animals,

specifically map on to disorders of social behaviour to complement imaging studies. Although lesion studies also have limitations such as smaller sample sizes and comorbid disorders they are important in advancing our knowledge of the function of neural regions.

7.4. Conclusions

The findings of this thesis suggest that: 1) specific components of empathy have distinct associations with different kinds of disrupted trait social-cognitive ability 2) specific components of empathy are positively associated with trait prosocial behaviour and individual differences in the capacity to regulate one's own emotions moderates the strength with which empathy relates to trait prosocial behaviour 3) anterior cingulate cortex function may be critical in the perception of vicarious information, including pain and reward processing; and individual differences in anterior cingulate cortex function during pain and reward processing relates to individual differences in empathy 4) empathy is critical for prosocial decision making and underpins the neural computations that signal outcomes for others that are different from our expectations. Together, these findings contribute to a more complete and coherent understanding of the structure, function and neural basis of empathic/vicarious processing.

References

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, *15*, 1–16.

Ali, F., & Chamorro-Premuzic, T. (2010). Investigating Theory of Mind deficits in nonclinical psychopathy and Machiavellianism. *Personality and Individual Differences*, *49*, 169–174.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews. Neuroscience*, *7*, 268–277.

Anastassiou-Hadjicharalambous, X., & Warden, D. (2008). Physiologically-indexed and self-perceived affective empathy in conduct-disordered children high and low on callous-unemotional traits. *Child Psychiatry and Human Development*, *39*, 503–517.

Anderson, N. E., & Kiehl, K. A. (2012). The psychopath magnetized: insights from brain imaging. *Trends in Cognitive Sciences*, *16*, 52–60.

Apps, M. A. J., Green, R., & Ramnani, N. (2013). Reinforcement learning signals in the anterior cingulate cortex code for others' false beliefs. *NeuroImage*, *64*, 1–9.

Apps, M. A. J., Lesage, E., & Ramnani, N. (2015). Vicarious reinforcement learning signals when instructing others. *The Journal of Neuroscience*, *35*, 2904–2913.

Apps, M. A. J., Lockwood, P. L., & Balsters, J. H. (2013). The role of the midcingulate cortex in monitoring others' decisions. *Frontiers in Neuroscience*, *7*, 251.

Apps, M. A. J., & Ramnani, N. (2014). The anterior cingulate gyrus signals the net value of others' rewards. *The Journal of Neuroscience*, *34*, 6190–6200.

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, *26*, 839–851.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*, 23–32.

Bailey, P. E., Ruffman, T., & Rendell, P. G. (2013). Age-Related Differences in Social Economic Decision Making : The Ultimatum Game, *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences 68*, 356–363.

Balsters, J., Apps, M. A. J., Bolis, D., Lehner, R., Gallagher, L., & Wenderoth, N. (2015). Deficits in the neural mechanisms underpinning social prediction errors in Autism Spectrum Disorder. In *Social functioning and autism*. Organisation for Human Brain Mapping, Hawaii convention centre.

Barbas, H., Ghashghaei, H., Dombrowski, S. M., & Rempel-Clower, N. L. (1999). Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *Journal of Comparative Neurology*, *410*, 343–367.

Baron-Cohen, S. (2011). *Zero degrees of empathy: A new theory of human cruelty*. Penguin UK.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning

autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*, 163–175.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*, 5–17.

Batson, C. D. (1998). Altruism and prosocial behavior. In *The handbook of social psychology, 4th edition*. New York: McGraw-Hill.

Batson, C. D. (2009). These things called empathy: eight related but distinct phenomena. In *The social neuroscience of empathy* (Decety, Jean, Ickes, William, pp. 3–15). Cambridge, MA, US: MIT Press.

Beadle, J. N., Sheehan, A. H., Dahlben, B., & Gutchess, A. H. (2013). Aging, Empathy, and Prosociality. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 1–10.

Beckmann, M., Johansen-Berg, H., & Rushworth, M. F. S. (2009). Connectivity-Based Parcellation of Human Cingulate Cortex and Its Relation to Functional Specialization. *Journal of Neuroscience*, *29*, 1175–1190.

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science*, *324*, 1160–1164.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*, 245–249.

Benedetti, F., Mayberg, H. S., Wager, T. D., Stohler, C. S., & Zubieta, J.-K. (2005). Neurobiological mechanisms of the placebo effect. *The Journal of Neuroscience*, *25*, 10390–10402.

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

Berman, A. H., Bergman, H., Palmstierna, T., & Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *European Addiction Research*, *11*, 22–31.

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience*, *35*, 1–23.

Bernhardt, B. C., Valk, S. L., Silani, G., Bird, G., Frith, U., & Singer, T. (2014). Selective disruption of sociocognitive structural brain networks in autism and alexithymia. *Cerebral Cortex*, 3258–3267.

Bird, G., & Cook, R. (2013). Mixed emotions : The contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, e285.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, *133*, 1515–1525.

Bird, G., & Viding, E. (2014). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520–532.

Blair. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, *14*, 698–718.

Blair, J., Mitchell, D., & Blair, K. (2005). *The Psychopath: emotion and the Brain.* Blackwell Publishing.

Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., & Jones, L. (1996). Theory of mind in the psychopath. *Journal of Forensic Psychiatry*, *7*, 15–25.

Blair, R. (1999). Psychophysiological responsiveness to the distress ofothers in children with autism. *Personality and Individual Differences*, *26*, 477–485.

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews Neuroscience*, 786–799.

Blair, R. J. R., Colledge, E., Murray, L., & Mitchell, D. G. V. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology*, *29*, 491–498.

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology*, *34*, 192–198.

Boorman, E. D., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The Behavioral and Neural Mechanisms Underlying the Tracking of Expertise. *Neuron*, *80*, 1558 – 1571.

Braams, B. R., Güroglu, B., de Water, E., Meuwese, R., Koolschijn, P. C., Peper, J. S., & Crone, E. A. (2014). Reward-related neural responses are dependent on the beneficiary. *Social Cognitive and Affective Neuroscience*, *9*, 1030–1037.

Brazil, I. A., Mars, R. B., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., & de Bruijn, E. R. a. (2011). A neurophysiological dissociation between monitoring one's own and others' actions in psychopathy. *Biological Psychiatry*, *69*, 693–699.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage*, *16*, S497.

Brook, M., Brieman, C. L., & Kosson, D. S. (2013). Emotion processing in Psychopathy Checklist—assessed psychopathy: A review of the literature. *Clinical Psychology Review*, *33*, 979–995.

Brook, M., & Kosson, D. S. (2013). Impaired cognitive empathy in criminal psychopathy: evidence from a laboratory measure of empathic accuracy. *Journal of Abnormal Psychology*, *122*, 156–166.

Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., & Boussaoud, D. (2008). Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory. *Cerebral Cortex*, *18*, 1485–1495.

Brugha, T. S., McManus, S., Smith, J., Scott, F., Meltzer, H., Purdon, S., … Radley, J. (2012). Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. *Psychological Medicine*, *42*, 647–656.

Buchman, A. S., Boyle, P. A., Wilson, R. S., Fleischman, D. A., Leurgans, S., & Bennett, D. A. (2009). Association between Late-Life Social Activity and Motor Decline in Older Adults. *Archives of Internal Medicine*, *169*, 1139–1146.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14431–14436.

Capps, L., Kasari, C., Yirmiya, N., & Sigman, M. (1993). Parental perception of emotional expressiveness in children with autism. *Journal of Consulting and Clinical Psychology*, *61*, 475–484.

Carlo, G., Allen, J. B., & Buhman, D. C. (1999). Facilitating and disinhibiting prosocial behaviors: The nonlinear interaction of trait perspective taking and trait personal distress on volunteering. *Basic and Applied Social Psychology*, *21*, 189–197.

Carlo, G., Hausmann, A., Christiansen, S., & Randall, B. A. (2003). Sociocognitive and Behavioral Correlates of a Measure of Prosocial Tendencies for Adolescents. *Journal of Early Adolescence*, *23*, 107–134.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*, 1839–1849.

Chakraborty, A., & Chakrabarti, B. (2015). Is it me? Self-recognition bias across sensory modalities and its relationship to autistic traits. *Molecular Autism*, *6*, 20.

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

Chang, S. W. C., Gariépy, J.-F., & Platt, M. L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neuroscience*, *16*, 243–250.

Cheng, Y., Hung, A.-Y., & Decety, J. (2012). Dissociation between affective sharing and emotion understanding in juvenile psychopaths. *Development and Psychopathology*, *24*, 623–636.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*, 231–239.

Cleckley, H. (1941). The mask of sanity; an attempt to reinterpret the so-called psychopathic personality.

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: from origin to function. *Behavioral and Brain Sciences*, *37*, 177–192.

Cooper, J. C., Dunne, S., Furey, T., & O'Doherty, J. P. (2012). Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. *Journal of Cognitive Neuroscience*, *24*, 106–118.

Cope, L. M., Shane, M. S., Segall, J. M., Nyalakanti, P. K., Stevens, M. C., Pearlson, G. D., … Kiehl, K. A. (2012). Examining the effect of psychopathic traits on gray matter volume in a community substance abuse sample. *Psychiatry Research-Neuroimaging*, *204*, 91–100.

Cracknell, R. (2010). The ageing population, 2010.

Craig, A. D. (2009). How do you feel - now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*, 59–70.

Curry, O., Chesters, M. J., & Viding, E. (2011). The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. *Personality and Individual Differences*, *50*, 804–809.

Davis, M. H. (1983). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality*, *51*, 167–184.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision Making, Affect, and Learning: Attention and Performance XXIII*, *23*, 1.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology*, *40*, 271.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology*, *27*, 403–424.

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*, 285–298.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive Affective & Behavioral Neuroscience*, *8*, 429–453.

De Brito, S. A., Mechelli, A., Wilke, M., Laurens, K. R., Jones, A. P., Barker, G. J., … Viding, E. (2009). Size matters: Increased grey matter in boys with conduct problems and callousunemotional traits. *Brain*, *132*, 843–852.

Decety, J., & Ickes, W. (2011). *The social neuroscience of empathy*. MIT Press.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 71–100.

Decety, J., Michalska, K. J., Akitsuki, Y., & Lahey, B. B. (2009). Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. *Biological Psychology*, *80*, 203–211.

Decety, J., Skelly, L. R., & Kiehl, K. A. (2013). Brain response to empathy-eliciting scenarios involving pain in incarcerated individuals with psychopathy. *JAMA Psychiatry*, *70*, 638–645.

Delmonte, S., Gallagher, L., O'Hanlon, E., McGrath, J., & Balsters, J. H. (2013). Functional and structural connectivity of frontostriatal circuitry in Autism Spectrum Disorder. *Frontiers in Human Neuroscience*, *430*.

Depue, R. A., & Morrone-Strupinsky, J. V. (2005). A neurobehavioral model of affiliative bonding: Implications for conceptualizing a human trait of affiliation. *Behavioral and Brain Sciences*, *28*, 313–349.

de Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences*, *10*, 435–41.

De Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.*, *59*, 279–300.

Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine*, *34*, 1093–1102.

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, *38*, 464–473.

Eisenberger, N. I., Berkman, E. T., Inagaki, T. K., Rameson, L. T., Mashal, N. M., & Irwin, M. R. (2010). Inflammation-induced anhedonia: endotoxin reduces ventral striatum responses to reward. *Biological Psychiatry*, *68*, 748–754.

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, *51*, 665–697.

Eisenberg, N., Eggum, N. D., & Di Giunta, L. (2010). Empathy-related Responding: Associations with Prosocial Behavior, Aggression, and Intergroup Relations. *Social Issues and Policy Review*, *4*, 143–180.

Eisenberg, N., & Fabes, R. A. (1992). Emotion, regulation, and the development of social competence. In *Emotion and social behavior* (pp. 119–150). Thousand Oaks, CA, US: Sage Publications, Inc.

Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: a multimethod study. *Journal of Personality and Social Psychology*, *57*, 55–66.

Eisenberg, N., McCreath, H., & Ahn, R. (1988). Vicarious Emotional Responsiveness and Prosocial Behavior Their Interrelations in Young Children. *Personality and Social Psychology Bulletin*, *14*, 298–311.

Engen, H. G., & Singer, T. (2013). Empathy circuits. *Current Opinion in Neurobiology*, *23*, 275–282.

Ertel, K. a, Glymour, M. M., & Berkman, L. F. (2008). Effects of social integration on preserving memory function in a nationally representative US elderly population. *American Journal of Public Health*, *98*, 1215–20.

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous-unemotional traits in a community sample of adolescents. *Assessment*, *13*, 454–469.

Eysenck, S. B., & Eysenck, H. J. (1978). Impulsiveness and venturesomeness: Their position in a dimensional system of personality description. *Psychological Reports*, *43*, 1247–1255.

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews*, *35*, 903–911.

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, *46*, 160–173.

Fan, Y.-T., Chen, C., Chen, S.-C., Decety, J., & Cheng, Y. (2014). Empathic arousal and social understanding in individuals with autism: evidence

from fMRI and ERP measurements. *Social Cognitive and Affective Neuroscience*, *9*, 1203–1213.

Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., & Delgado, M. R. (2012). Social network modulation of reward-related signals. *The Journal of Neuroscience*, *32*, 9045–9052.

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences*, *11*, 419–427.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785–91.

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, *8*, 185–190.

Foulkes, L., Bird, G., Gökçen, E., McCrory, E., & Viding, E. (2015). Common and Distinct Impacts of Autistic Traits and Alexithymia on Social Reward. *PloS One*, *10*.

Foulkes, L., McCrory, E. J., Neumann, C. S., & Viding, E. (2014). Inverted Social Reward: Associations between Psychopathic Traits and Self-Report and Experimental Measures of Social Reward. *PLoS ONE*, *9*, e106000.

Fratiglioni, L., Paillard-Borg, S., & Winblad, B. (2004). An active and socially integrated lifestyle in late life might protect against dementia. *The Lancet Neurology*, *4*, 343–353.

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Annual Research Review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. *Journal of Child Psychology and Psychiatry*, *55*, 532–548.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*, 531–534.

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287–313.

Fumagalli, M., & Priori, A. (2012). Functional and clinical neuroanatomy of morality. *Brain*, *135*, 2006–2021.

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

Gadow, K. D., & Sprafkin, J. (2009). *The Symptom Inventories: An annotated bibliography.* Stony Brook, NY: Checkmate Plus.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*, 581–586.

Gordon, R. (1986). Folk psychology as simulation. *Mind and Language*, *1*, 158–171.

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*, 224–237.

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*, 359–365.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*, 348–362.

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, *64*, 970.

Gross, J. J., & Thompson, R. A. (2007). Emotion Regulation: Conceptual Foundations. In *Handbook of emotion regulation* (pp. 3–24). New York, NY, US: Guilford Press.

Gu, X., Liu, X., Guise, K. G., Naidich, T. P., Hof, P. R., & Fan, J. (2010). Functional dissociation of the frontoinsular and anterior cingulate cortices in empathy for pain. *The Journal of Neuroscience*, *30*, 3739–3744.

Haber, S. N., Kunishio, K., Mizobuchi, M., & Lynd-Balta, E. (1995). The orbital and medial prefrontal circuit through the primate basal ganglia. *The Journal of Neuroscience*, *15*, 4851–4867.

Hadland, K. A., Rushworth, M. F. S., Gaffan, D., & Passingham, R. E. (2003). The effect of cingulate lesions on social behaviour and emotion. *Neuropsychologia*, *41*, 919–931.

Haffey, A., Press, C., O'Connell, G., & Chakrabarti, B. (2013). Autistic Traits Modulate Mimicry of Social but not Nonsocial Rewards. *Autism Research*, *6*, 614–620.

Hampton, A. N., Bossaerts, P., & Doherty, J. P. O. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans.

Hansen, A. L., Johnsen, B. H., Hart, S., Waage, L., & Thayer, J. F. (2008). Brief communication: Psychopathy and recognition of facial expressions of emotion. *Journal of Personality Disorders*, *22*, 639–644.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*, *316*, 1622–1625.

Hare, R. D. (1999). *The Hare Psychopathy Checklist-Revised: PLC-R*. MHS, Multi-Health Systems.

Hare, R. D., & Neumann, C. N. (2006). The PCL-R Assessment of Psychopathy: Development, Structural Properties, and New Directions. In C. Patrick (Ed.), *Handbook of Psychopathy* (Patrick, C, pp. 58–88). New York: Guilford.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a Clinical and Empirical Construct. *Annual Review of Clinical Psychology*, *4*, 217–246.

Haroush, K., & Williams, Z. M. (2015). Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates. *Cell*, 1–13.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science*, *324*, 948–950.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Heal, J. (1996). Simulation, theory, and content. *Theories of Theories of Mind*, 75–89.

Hickok, G. (2014). *The myth of mirror neurons: the real neuroscience of communication and cognition*. WW Norton & Company.

Hicks, B., & Patrick, C. (2006). Psychopathy and Negative Emotionality: Analyses of Suppressor Effects Reveal Distinct Relations With Emotional Distress, Fearfulness, and Anger–Hostility. *Journal of Abnormal Psychology*, *115*, 276–287.

Hill, E. L., & Frith, U. (2003). Understanding autism: insights from mind and brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*, 281–289.

Hillman, K. L., & Bilkey, D. K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. *Nature Neuroscience*, *15*, 1290–1297.

Hoffman, M. L. (2001). *Empathy and moral development: Implications for caring and justice*. Cambridge University Press.

Hoffman, M. L. (2008). Empathy and prosocial behavior. *Handbook of Emotions*, *3*, 440–455.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, *33*, 307.

Hsu, M., Anen, C., & Quartz, S. R. (2008). The right and the good: distributive justice and neural encoding of equity and efficiency. *Science*, *320*, 1092–1095.

Hudry, K., & Slaughter, V. (2009). Agent familiarity and emotional context influence the everyday empathic responding of young children with autism. *Research in Autism Spectrum Disorders*, *3*, 74–85.

Hunt, L., & Behrens, T. (2011). Frames of reference in human social decision making. In *Neural basis of motivational and cognitive control*. MIT Press.

Iannetti, G. D., & Mouraux, A. (2011). Can the functional MRI responses to physical pain really tell us why social rejection 'hurts'? *Proceedings of the National Academy of Sciences of the United States of America*, *108*, E343; author reply E344.

Insel, T. R., & Young, L. J. (2001). The neurobiology of attachment. *Nature Reviews Neuroscience*, *2*, 129–136.

Jabbi, M., Bastiaansen, J., & Keysers, C. (2008). A Common Anterior Insula Representation of Disgust Observation, Experience and Imagination Shows Divergent Functional Connectivity Pathways. *Plos One*, *3*, 8.

James, B. D., Wilson, R. S., Barnes, L. L., & Bennett, D. a. (2011). Late-life social activity and cognitive decline in old age. *Journal of the International Neuropsychological Society : JINS*, *17*, 998–1005.

Johansen-Berg, H., Gutman, D., Behrens, T., Matthews, P., Rushworth, M. F. S., Katz, E., … Mayberg, H. (2008). Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression. *Cerebral Cortex*, *18*, 1374–1383.

John, O. P., & Gross, J. J. (2004). Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development. *Journal of Personality*, *72*, 1301–1334.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*, 589–611.

Jones, A. P., Happé, F. G. E., Gilbert, F., Burnett, S., & Viding, E. (2010). Feeling, caring, knowing: different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *11*, 1188–1197.

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., … Casey, B. J. (2011). Behavioral and neural properties of social reinforcement learning. *The Journal of Neuroscience*, *31*, 13039–45.

Kemp, J., Després, O., Sellal, F., & Dufour, A. (2012). Theory of Mind in normal ageing and neurodegenerative pathologies. *Ageing Research Reviews*, *11*, 199–219.

Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in Cognitive Sciences*, *16*, 559–572.

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. a, Cruise, K., Munoz, L. C., … Morris, A. S. (2008). Assessing callous-unemotional traits in adolescent offenders: validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry*, *31*, 241–252.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, *314*, 829–832.

Koenigs, M., Kruepke, M., & Newman, J. (2010). Economic decision-making in psychopathy: A comparison with ventromedial prefrontal lesion patients. *Neuropsychologia*, *48*, 2198–2204.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*, 2492–2502.

Lamm, C., Silani, G., & Singer, T. (2015). Distinct neural networks underlying empathy for pleasant and unpleasant touch. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *70*, 79–89.

Lander, G. C., Lutz-Zois, C. J., Rye, M. S., & Goodnight, J. A. (2012). The differential association between alexithymia and primary versus secondary psychopathy. *Personality and Individual Differences*, *52*, 45–50.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nature Neuroscience*, *11*, 404–409.

Levenston, G. K., Patrick, C. J., Bradley, M. M., & Lang, P. J. (2000). The psychopath as observer: emotion and attention in picture processing. *Journal of Abnormal Psychology*, *109*, 373–385.

Lilienfeld, S. O., Fowler, K. A., & Patrick, C. (2006). The self-report assessment of psychopathy. In *Handbook of psychopathy* (pp. 107–132). Guildford press.

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task.

Proceedings of the National Academy of Sciences of the United States of America, *104*, 9493–9498.

Louth, S., Hare, R., & Linden, W. (1998). Psychopathy and alexithymia in female offenders. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *30*, 91–98.

Ly, M., Motzkin, J. C., Philippi, C. L., Kirk, G. R., Newman, J. P., Kiehl, K. A., & Koenigs, M. (2012). Cortical Thinning in Psychopathy. *American Journal of Psychiatry*, *169*, 743–749.

Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*, *116*, 155–65.

Lynd-Balta, E., & Haber, S. N. (1994). Primate striatonigral projections: A comparison of the sensorimotor-related striatum and the ventral striatum. *The Journal of Comparative Neurology*, *345*, 562–578.

Markowitsch, H. J., Emmans, D., Irle, E., Streicher, M., & Preilowski, B. (1985). Cortical and subcortical afferent connections of the primate's temporal pole: A study of rhesus monkeys, squirrel monkeys, and marmosets. *The Journal of Comparative Neurology*, *242*, 425–458.

Masten, C. L., Colich, N. L., Rudie, J. D., Bookheimer, S. Y., Eisenberger, N. I., & Dapretto, M. (2011). An fMRI investigation of responses to peer rejection in adolescents with autism spectrum disorders. *Developmental Cognitive Neuroscience*, *1*, 260–270.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., … Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, *45*, 651–660.

Medford, N., & Critchley, H. D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Structure and Function*, *214*, 535–549.

Meffert, H., Gazzola, V., Boer, J. A. den, Bartels, A. A. J., & Keysers, C. (2013). Reduced spontaneous but relatively normal deliberate vicarious representations in psychopathy. *Brain*, *136*, 2550–2562.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*, 525–543.

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., … Dalgleish, T. (2009). A Key Role for Similarity in Vicarious Reward. *Science*, *324*, 900–900.

Mokros, A., Menner, B., Eisenbarth, H., Alpers, G. W., Lange, K. W., & Osterheider, M. (2008). Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *Journal of Abnormal Psychology*, *117*, 406–413.

Molenberghs, P., Bosworth, R., Nott, Z., Louis, W. R., Smith, J. R., Amiot, C. E., … Decety, J. (2014). The influence of group membership and individual differences in psychopathy and perspective taking on neural responses when punishing and rewarding others. *Human Brain Mapping*, *35*, 4989–4999.

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*, *103*, 15623–15628.

Moll, J., & Schulkin, J. (2009). Social attachment and aversion in human moral cognition. *Neuroscience & Biobehavioral Reviews*, *33*, 456–465.

Moran, J. M. (2013). Lifespan development: the effects of typical aging on theory of mind. *Behavioural Brain Research*, *237*, 32–40.

Morelli, S. A., Sacchet, M. D., & Zaki, J. (2015). Common and distinct neural correlates of personal and vicarious reward: A quantitative meta-analysis. *NeuroImage*, *112*, 244–253.

Mufson, E. J., & Mesulam, M.-M. (1982). Insula of the old world monkey. II: Afferent cortical input and comments on the claustrum. *The Journal of Comparative Neurology*, *212*, 23–37.

Neumann, C. S., & Pardini, D. (2014). Factor structure and construct validity of the Self-Report Psychopathy (SRP) scale and the Youth Psychopathic Traits Inventory (YPI) in young men. *Journal of Personality Disorders*, *28*, 419–433.

Neumann, C. S., Schmitt, D. S., Carter, R., Embley, I., & Hare, R. D. (2012). Psychopathic Traits in Females and Males across the Globe. *Behavioral Sciences & the Law*, *30*, 557–574.

Oakley, B. A. (2013). Concepts and implications of altruism bias and pathological altruism. *Proceedings of the National Academy of Sciences of the United States of America*, *110 Suppl 2*, 10408–10415.

Palomero-Gallagher, N., Mohlberg, H., Zilles, K., & Vogt, B. (2008). Cytology and receptor architecture of human anterior cingulate cortex. *Journal of Comparative Neurology*, *508*, 906–926.

Panasiti, M. S., Puzzo, I., & Chakrabarti, B. (2015). Autistic Traits Moderate the Impact of Reward Learning on Social Behaviour. *Autism Research*, 1–9.

Pandya, D. N., Van Hoesen, G. W., & Mesulam, M. M. (1981). Efferent connections of the cingulate gyrus in the rhesus monkey. *Experimental Brain Research*, *42*, 319–330.

Paulhus, D., Neumann, C., & Hare, R. (2015). *Manual for the self-report psychopathy scale*. Toronto: Multi-Health systems.

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: multilevel perspectives. *Annual Review of Psychology*, *56*, 365–392.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*, 1042–1045.

Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology*, *20*, 51–63.

Rademacher, L., Krach, S., Kohls, G., Irmak, A., Gründer, G., & Spreckelmeyer, K. N. (2010). Dissociation of neural networks for anticipation and consumption of monetary and social rewards. *Neuroimage*, *49*, 3276–3285.

Ray, R. D., Ochsner, K. N., Cooper, J. C., Robertson, E. R., Gabrieli, J. D. E., & Gross, J. J. (2005). Individual differences in trait rumination and the neural systems supporting cognitive reappraisal. *Cognitive, Affective & Behavioral Neuroscience*, *5*, 156–168.

Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. a. (2011). The QCAE: a Questionnaire of Cognitive and Affective Empathy. *Journal of Personality Assessment*, *93*, 84–95.

Rescorla, R. A., & Wagner, A. R. (1972). Classical Conditioning II: Current Research and Theory. In W. F. Prokasy (Ed.), (pp. 64–99). New York: Appleton-Century Crofts.

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*, 410–424.

Richell, R., Mitchell, D., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. (2003). Theory of mind and psychopathy: can psychopathic individuals read the 'language of the eyes'? *Neuropsychologia*, *41*, 523–526.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, *62*, 23–48.

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happé, F., … Ronald, A. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, *68*, 1113–1121.

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., & Smith, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, *55*, 594–602.

Romeo, R., Knapp, M., & Scott, S. (2006). Economic cost of severe antisocial behaviour in children--and who pays it. *The British Journal of Psychiatry: The Journal of Mental Science*, *188*, 547–553.

Rudebeck, P. H., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2006). A role for the macaque anterior cingulate gyrus in social valuation. *Science*, *313*, 1310–1312.

Rudebeck, P. H., Putnam, P. T., Daniels, T. E., Yang, T., Mitz, A. R., Rhodes, S. E. V., & Murray, E. A. (2014). A role for primate subgenual cingulate

cortex in sustaining autonomic arousal. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 5391–5396.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews. Neuroscience*, *15*, 549–562.

Rushworth, M. F. S., Behrens, T. E. J., Rudebeck, P. H., & Walton, M. E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in Cognitive Sciences*, *11*, 168–176.

Rushworth, M. F. S., Mars, R. B., & Sallet, J. (2013). Are there specialized circuits for social cognition and are they unique to humans? *Current Opinion in Neurobiology*, *23*, 436–442.

Rushworth, M. F. S., Mars, R. B., & Summerfield, C. (2009). General mechanisms for making decisions? *Current Opinion in Neurobiology*, *19*, 75–83.

Rütgen, M., Seidel, E.-M., Riečanský, I., & Lamm, C. (2015). Reduction of Empathy for Pain by Placebo Analgesia Suggests Functional Equivalence of Empathy and First-Hand Emotion Experience. *The Journal of Neuroscience*, *35*, 8938–8947.

Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X., Jbabdi, S., … Rushworth, M. F. S. (2011). Social network size affects neural circuits in macaques. *Science*, *334*, 697–700.

Sallet, J., Mars, R. B., Noonan, M. P., Neubert, F.-X., Jbabdi, S., O'Reilly, J. X., … Rushworth, M. F. S. (2013). The organization of dorsal frontal cortex in humans and macaques. *The Journal of Neuroscience*, *33*, 12255–12274.

Sallet, J., Quilodran, R., Rothe, M., Vezoli, J., Joseph, J. P., & Procyk, E. (2007). Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive Affective & Behavioral Neuroscience*, *7*, 327–336.

Sanfey, A. G. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, *318*, 598–602.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*, 1755–1758.

Schipper, M., & Petermann, F. (2013). Relating empathy and emotion regulation: do deficits in empathy trigger emotion dysregulation? *Social Neuroscience*, *8*, 101–107.

Schülke, O., Bhagavatula, J., Vigilant, L., & Ostner, J. (2010). Social bonds enhance reproductive success in male macaques. *Current Biology*, *20*, 2207–2210.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115.

Schultz, W. (2013). Updating dopamine reward signals. *Current Opinion in Neurobiology*, *23*, 229 – 238.

Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., … Freitag, C. M. (2012). Empathy in children with autism and conduct disorder: group-specific profiles and developmental aspects. *Journal Of Child Psychology And Psychiatry*, *53*, 651–659.

Seara-Cardoso, A., Dolberg, H., Neumann, C., Roiser, J. P., & Viding, E. (2013). Empathy, morality and psychopathic traits in women. *Personality and Individual Differences*, *55*, 328–333.

Seara-cardoso, A., Neumann, C., Roiser, J., Mccrory, E., & Viding, E. (2012). Investigating associations between empathy , morality and psychopathic personality traits in the general population. *Personality and Individual Differences*, *52*, 67–71.

Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S.-J., Brito, S. A. D., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*, 53–63.

Sebastian, C. L., McCrory, E. J. P., Cecil, C. A. M., Lockwood, P. L., De Brito, S. A., Fontaine, N. M. G., & Viding, E. (2012). Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. *Archives of General Psychiatry*, *69*, 814–822.

Seltzer, B., & Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *The Journal of Comparative Neurology*, *281*, 97–113.

Shidara, M., & Richmond, B. J. (2002). Anterior cingulate: Single neuronal signals related to degree of reward expectancy. *Science*, *296*, 1709–1711.

Shultz, S., & Dunbar, R. I. M. (2010). Social bonds in birds are associated with brain size and contingent on the correlated evolution of life-history and increased parental investment. *Biological Journal of the Linnean Society*, *100*, 111–123.

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Social Neuroscience*, *3*, 97–112.

Silk, J. B. (2007). Social components of fitness in primate groups. *Science*, *317*, 1347–1351.

Simms, M. L., Kemper, T. L., Timbie, C. M., Bauman, M. L., & Blatt, G. J. (2009). The anterior cingulate cortex in autism: heterogeneity of qualitative and quantitative cytoarchitectonic features suggests possible subgroups. *Acta Neuropathologica*, *118*, 673–684.

Singer, & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, *1156*, 81–96.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews*, *30*, 855–863.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, *24*, R875–R878.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*, 1157–1162.

Smith, A. (2009). Emotional empathy in autism spectrum conditions: weak, intact, or heightened? *Journal of Autism and Developmental Disorders*, *39*, 1747–8; author reply 1749–54.

Soto, J. A., Perez, C. R., Kim, Y.-H., Lee, E. A., & Minnick, M. R. (2011). Is expressive suppression always associated with poorer psychological functioning? A cross-cultural comparison between European Americans and Hong Kong Chinese. *Emotion*, *11*, 1450–1455.

Sterzer, P., Stadler, C., Krebs, A., Kleinschmidt, A., & Poustka, F. (2005). Abnormal neural responses to emotional visual stimuli in adolescents with conduct disorder. *Biological Psychiatry*, *57*, 7–15.

Sterzer, P., Stadler, C., Poustka, F., & Kleinschmidt, A. (2007). A structural neural deficit in adolescents with conduct disorder and its association with lack of empathy. *NeuroImage*, *37*, 335–42.

Sullivan, S., & Ruffman, T. (2004). Social understanding: How does it fare with advancing years? *British Journal of Psychology*, *95*, 1–18.

Sul, S., Tobler, P. N., Hein, G., Leiberg, S., Jung, D., Fehr, E., & Kim, H. (2015). Spatial gradient in value representation along the medial prefrontal cortex reflects individual differences in prosociality. *Proceedings of the National Academy of Sciences*, *112*, 7851–7856.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: an introduction*. Cambridge, Massachusetts: MIT press.

Swart, M., Kortekaas, R., & Aleman, A. (2009). Dealing with feelings: characterization of trait alexithymia on emotion regulation strategies and cognitive-emotional processing. *PloS One*, *4*, e5751.

Sze, J. A., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (2012). Greater Emotional Empathy and Prosocial Behavior in Late Life. *Emotion (Washington, D.C.)*, *12*, 1129–1140.

Szolnoki, A., Xie, N.-G., Wang, C., & Perc, M. (2011). Imitating emotions instead of strategies in spatial games elevates social welfare. *EPL (Europhysics Letters)*, *96*, 38002.

Szolnoki, A., Xie, N.-G., Ye, Y., & Perc, M. (2013). Evolution of emotions on networks leads to the evolution of cooperation in social dilemmas. *Physical Review E*, *87*, 042805.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of

activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289.

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A. R. B., & Crone, E. A. (2009). What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Social Cognitive and Affective Neuroscience*, *4*, 294–304.

Viding, E., Blair, R. J. R., Moffitt, T. E., & Plomin, R. (2005). Evidence for substantial genetic risk for psychopathy in 7-year-olds. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *46*, 592–597.

Viding, E., McCrory, E., & Seara-Cardoso, A. (2014). Psychopathy. *Current Biology: CB*, *24*, R871–874.

Vogt, B. A., Nimchinsky, E. A., Vogt, L. J., & Hof, P. R. (1995). Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. *The Journal of Comparative Neurology*, *359*, 490–506.

Watson, K. K., & Platt, M. L. (2012). Social signals in primate orbitofrontal cortex. *Current Biology*, *22*, 2268–2273.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, Texas: Harcourt Assessment.

Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. *NeuroImage*, *33*, 493–504.

White, B. a. (2014). Who cares when nobody is watching? Psychopathic traits and empathy in prosocial behaviors. *Personality and Individual Differences*, *56*, 116–121.

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*, 655–664.

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*, 364–372.

Williams, S. M., & Goldman-Rakic, P. S. (1998). Widespread origin of the primate mesofrontal dopamine system. *Cerebral Cortex*, *8*, 321–345.

Willis, M. L., Palermo, R., Burke, D., McGrillen, K., & Miller, L. (2010). Orbitofrontal cortex lesions result in abnormal social judgements to emotional faces. *Neuropsychologia*, *48*, 2182–2187.

Woo, C.-W., Koban, L., Kross, E., Lindquist, M. A., Banich, M. T., Ruzic, L., … Wager, T. D. (2014). Separate neural representations for physical pain and social rejection. *Nature Communications*, *5*, 5380.

Yeterian, E. H., & Pandya, D. N. (1991). Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. *The Journal of Comparative Neurology*, *312*, 43–67.

Yoder, K. J., & Decety, J. (2014). Spatiotemporal neural dynamics of moral judgment: a high-density ERP study. *Neuropsychologia*, *60*, 39–45.

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, *7*, 1–10.

Zahn, R., de Oliveira-Souza, R., Bramati, I., Garrido, G., & Moll, J. (2009). Subgenual cingulate activity reflects individual differences in empathic concern. *Neuroscience Letters*, *457*, 107–110.

Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., & Grafman, J. (2009). The neural basis of human social values: evidence from functional MRI. *Cerebral Cortex*, *19*, 276–283.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature Neuroscience*, *15*, 675–680.

Appendix 1: QCAE - Questionnaire of Cognitive and Affective Empathy

Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: a Questionnaire of Cognitive and Affective Empathy. Journal of Personality Assessment, 93(1), 84–95.

1. I sometimes find it difficult to see things from the "other guy's" point of view (**OS r**)

2. I am usually objective when I watch a film or play, and I don't often get completely caught up in it (**PeR r**)

3. I try to look at everybody's side of a disagreement before I make a decision (**OS**)

4. I sometimes try to understand my friends better by imagining how things look from their perspective (**OS**)

5. When I am upset at someone, I usually try to "put myself in his shoes" for a while. (**OS**)

6. Before criticizing somebody, I try to imagine how I would feel if I was in their place. (**OS**)

7. I often get emotionally involved with my friends' problems. (**PrR**)

8. I am inclined to get nervous when others around me seem to be nervous. (**EC**)

9. People I am with have a strong influence on my mood. (**EC**)

10. It affects me very much when one of my friends seems upset. (**PrR**)

11. I often get deeply involved with the feelings of a character in a film, play, or novel. (**PeR**)

12. I get very upset when I see someone cry. (**PrR**)

13. I am happy when I am with a cheerful group and sad when the others are glum. (**EC**)

14. It worries me when others are worrying and panicky. (**EC**)

15. I can easily tell if someone else wants to enter a conversation. (**PT**)

16. I can pick up quickly if someone says one thing but means another. (**PT**)

17. It is hard for me to see why some things upset people so much. (**PeR r**)

18. I find it easy to put myself in somebody else's shoes. (**OS**)

19. I am good at predicting how someone will feel. (**PT**)

20. I am quick to spot when someone in a group is feeling awkward or uncomfortable. (**PT**)

21. Other people tell me I am good at understanding how they are feeling and what they are (**PT**) thinking.

22. I can easily tell if someone else is interested or bored with what I am saying. (**PT**)

23. Friends talk to me about their problems as they say that I am very understanding. (**PrR**)

24. I can sense if I am intruding, even if the other person does not tell me. (**PT**)

25. I can easily work out what another person might want to talk about. (**PT**)

26. I can tell if someone is masking their true emotion. (**PT**)

27. I am good at predicting what someone will do. (**PT**)

28. I can usually appreciate the other person's viewpoint, even if I do not agree with it. (**OS**)

29. I usually stay emotionally detached when watching a film. (**PeR r**)

30. I always try to consider the other fellow's feelings before I do something. (**OS**)

31. Before I do something I try to consider how my friends will react to it. (**OS**)

Cognitive Empathy

Perspective Taking (PT)

Online Simulation (OS)

Affective Empathy

Emotion Contagion (EC)

Proximal Responsivity (PrR)

Peripheral Responsivity (PeR)

(r – indicates that the item is reverse scored

Appendix 2: Main effects and group x condition interaction

Brain Region	BA	L/R	Peak voxel			k	t	z	FWE-corrected p-value
Pain>No pain									
Middle temporal gyrus	37	L	-48	-62	0	3817	14.83	>8.00	<.001
ext. middle occipital gyrus	19	L	-34	-86	-2		9.99	7.46	
ext. fusiform gyrus/inferior temporal gyrus	37	L	-46	-48	-16		9.41	7.18	
Secondary somatosensory cortex	40	L	-60	-30	34	1817	13.46	>8.00	<.001
ext. primary somatosensory cortex	2	L	-52	-26	36		12.44	>8.00	
Secondary somatosensory cortex	40	L	-60	-30	24		11.22	>8.00	
Inferior temporal gyrus	19	R	46	-56	-8	3739	13.22	>8.00	
ext. middle temporal gyrus	39	R	52	-62	-6		13.2	>8.00	<.001
ext. middle occipital gyrus	19	R	36	-80	8		10.93	>8.00	
IFG/DLPFC	9	L	-52	6	34	1433	10.09	7.51	<.001
ext. insula (middle)	13	L	-36	-4	14		5.48	4.86	
Primary somatosensory cortex	2	R	62	-22	30	1385	9.29	7.12	<.001
ext. secondary somatosensory cortex	40	R	36	-34	38		5.29	4.72	
DLPFC	46	R	52	40	4	2137	7.95	6.42	<.001
ext. frontal operculum	44/9	R	52	8	26		7.22	6.00	
IFG (triangularis)	45	L	-50	38	10	595	6.31	5.43	.002
Supplementary motor area	6	L	-30	-6	48	167	6.12	5.30	.003
Uncus/fusiform gyrus	20	L	-32	-2	-38	43	5.36	4.77	.03
No pain>Pain									
Middle temporal gyrus	20	L	-58	-34	-10	414	5.70	5.01	.01
Caudate nucleus	-	R	14	24	4	795	5.69	5.00	.01
Cerebellum	-	L	-42	-66	-38	135	5.60	4.94	.01
Lingual gyrus	18	R	12	-74	-4	802	5.37	4.78	.03
Group x Pain condition									**p-value (peak uncorrected)**
TD>CP (for Pain>No Pain)									
STG/insula (posterior)	22/13	L	-44	-16	-6	91	4.36	4.01	<.001
Cerebellum	-	R	0	-48	-20	74	3.82	3.57	<.001
Middle temporal gyrus	21	R	44	6	-36	24	3.42	3.23	.001
Caudate	-	R	10	10	10	22	3.34	3.17	.001
Supplementary motor area	6	L	-10	12	52	29	3.32	3.15	.001
IFG (orbitalis)	47	L	-22	10	-20	116	3.32	3.15	.001
ext. insula (anterior)	13	L	-32	12	-16		3.05	2.91	.002
Middle frontal gyrus *ext. IFG*	10	L	-36	34	20	266	3.23	3.07	.001
Thalamus	-	L	-4	-14	20	19	3.22	3.06	.001
Cerebellum	-	R	28	-58	-30	50	3.22	3.06	.001
Insula (posterior)	13	L	-28	-30	14	16	3.22	3.06	.001
Cerebellum	-	R	14	-72	-32	30	3.18	3.03	.001
SFG	10	R	22	48	22	38	3.17	3.02	.001
Globus pallidus	-	L	-16	-8	-4	18	3.15	3.00	.001
Brainstem/substantia nigra	-	R	10	-20	-16	25	3.07	2.93	.002
IFG (triangularis)	47	R	54	38	0	24	3.02	2.89	.002
Insula (anterior)	13	L	-30	16	2	29	3.00	2.87	.002
Anterior cingulate	24	L	0	20	24	14	2.92	2.80	.003
Precuneus	31	L	-4	-62	28	13	2.82	2.71	.003
CP>TD (for Pain>No Pain)									
Parahippocampal gyrus	-	L	-18	-20	-16	15	3.22	3.06	.001
Cerebellum	-	L	-2	-48	-4	17	3.08	2.94	.002

Main effects thresholded at $P<.05$ FWE-corrected at the peak level. Group x Condition interaction thresholded at $P<.005$ uncorrected, $k \geq 10$. BA= Putative Brodmann area; L/R=Left/Right; k=cluster size in $2mm^3$ voxels; DLPFC=Dorsolateral prefrontal cortex; IFG=Inferior frontal gyrus; SFG = Superior frontal gyrus; STG=Superior temporal gyrus; ext.=cluster extends into additional region.

Appendix 3: Uncorrected table of neural response at the time of the cue

Brain Region	BA	L/R	Peak voxel			k	t	z
Agency x Probability interaction (positive)								
Anterior cingulate*	24	R	8	32	12	340	6.47	5.05
Subcallosal gyrus		R	18	10	-15	123	4.79	4.08
		R	15	9	-6		3.79	3.39
Subcallosal gyrus		L	-12	3	-14	101	4.71	4.03
Occipital gyrus	37	R	36	63	7	44	4.53	3.91
Retrosplenial cortex	29	R	2	-34	21	60	4.1	3.61
Hippocampus		L	-15	-39	3	36	4.09	3.61
Caudate		L	-16	14	12	29	3.84	3.42
Superior temporal gyrus	41	R	46	-34	15	20	3.83	3.42
Caudate head		L	-14	40	6	12	3.79	3.39
Brainstem		R	9	-24	-14	18	3.75	3.36
Medial superior frontal gyrus	9	L	-6	45	18	15	3.68	3.31
Frontal pole	10	R	38	39	13	15	3.55	3.21
Agency x Probability interaction (negative)								
Superior temporal gyrus	38	R	30	9	-36	76	4.68	4.01
Other>Self								
Temporal pole*	38	R	33	22	-26	488	6.09	4.85
		R	50	21	-18		3.89	3.46
Insula	13	R	30	-25	21	52	4.85	4.12
Superior temporal gyrus	38	L	-50	-1	-12	56	4.47	3.87
Middle temporal gyrus	21	L	-60	3	23	55	4.46	3.86
Middle frontal gyrus	11	R	20	24	-9	27	4.45	3.85
Inferor Temporal Gyrus	20	L	-45	-12	-27	107	4.35	3.78
Uncus		L	-30	3	-32	22	3.96	3.51
Putamen		R	22	-6	10	19	3.96	3.51
Medial frontal gyrus	9	R	9	50	40	40	3.86	3.44
Temporal pole	38	L	-40	14	-33	23	3.71	3.33
Superior temporal gyrus	41	L	-46	-27	9	21	3.69	3.31
Posterior cingulate	31	R	12	-49	31	31	3.69	3.31
Middle frontal gyrus	8	R	30	24	51	12	3.67	3.3
Precuneus	31	L	-6	-69	22	27	3.66	3.29
Self > Other								
no voxels								
High probability>Low probability								
Ventral anterior cingulate	24	L	-3	27	1	29	3.85	3.44
Low probability >High probability								
Premotor cortex	6	R	24	-1	52	433	5.83	4.71
		R	24	9	49		5.11	4.28
Postcentral gyrus	2	R	38	-28	40	58	4.46	3.86
Superior temporal gyrus	41	L	-40	-40	13	22	4.17	3.66
Anterior cingulate sulcus	24	L	-12	0	42	15	4.17	3.66
Inferior parietal lobule	40	R	-51	-36	37	50	3.94	3.5
Inferior frontal gyrus	9	L	-54	3	31	13	3.87	3.44
Premotor Cortex	6	L	-28	-6	51	50	3.66	3.29
		L	-33	0	54		3.55	3.21
Middle frontal gyrus	9	L	-34	39	28	16	3.57	3.22

Interactions and main effects thresholded at $P<.001$ K=10. *survives FWE whole brain correction at the peak level.
BA=Putative Brodmann area; L/R=Left/Right; k=cluster size in 1.5mm^3 voxels.

Appendix 4: Uncorrected table of neural response at the time of the outcome

Brain Region	BA	L/R	Peak voxel			k	t	z
Agency x Outcome interaction (positive)								
Superior frontal gyrus	8	R	30	9	-36	76	4.68	4.01
Agency x Outcome interaction (negative)								
Putamen		L	-21	14	12	61	5.41	4.46
Superior frontal gyrus	6	L	-16	20	57	84	4.6	3.95
Precentral gyrus	6	L	-45	-7	45	60	4.55	3.92
Anterior cingulate	24	L	-6	6	30	55	4.34	3.78
		L	-12	14	30		3.63	3.27
Brainstem		R	4	-13	-14	13	4.16	3.65
Superior frontal gyrus	9	L	-20	32	36	43	4.11	3.62
Anterior cingulate	24	-	0	21	13	16	4.09	3.6
Insula	13	R	40	-15	-11	69	4.07	3.59
Middle temporal gyrus	21	R	48	-49	3	15	4.03	3.56
Postcentral gyrus	3	L	-27	-25	43	18	3.94	3.5
		L	-34	-25	40		3.8	3.39
Inferior frontal gyrus	47	L	-20	35	-6	17	3.92	3.48
Middle frontal gyrus	8	L	-27	15	36	11	3.8	3.39
Other outcome>Self outcome								
Caudate		R	9	16	18	72	4.65	3.98
Insula	13	L	-38	-6	-6	189	4.48	3.88
		L	-38	8	-8		4.46	3.86
Insula	13	L	-39	-6	25	36	4.48	3.88
Superior temporal gyrus	41	L	-40	-42	9	35	4.37	3.8
Medial frontal gyrus	11	R	4	60	-14	104	4.24	3.71
Posterior cingulate	29	R	3	-39	16	22	4.13	3.63
Cerebellum		R	28	-55	-54	31	4.05	3.58
Parrahippocampul gyrus	37	R	32	-40	-12	53	3.94	3.5
Posterior cingulate cortex	29	R	3	-46	21	60	3.89	3.46
Fusiform gyrus	37	R	33	-51	-14	11	3.81	3.4
Inferior occipital gyrus		R	40	-70	-5	29	3.81	3.4
		R	40	-75	-12		3.54	3.2
Insula	13	R	38	6	-11	21	3.76	3.36
Self outcome > Other outcome								
no voxels								
Win>No win								
Supplementary motor area	6		-6	2	57	13	3.94	3.5
No win> win								
no voxels								

Interactions and main effects thresholded at *P*<.001 K=10.BA=Putative Brodmann area; L/R=Left/Right; k=cluster size in 1.5mm^3 voxels.

Appendix 5: Published papers

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 November 2013
doi: 10.3389/fnhum.2013.00760

Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains

Patricia L. Lockwood[1]*, Geoffrey Bird[2,3], Madeleine Bridge[1] and Essi Viding[1]

[1] *Division of Psychology and Language Sciences, University College London, London, UK*
[2] *MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Kings College London, London, UK*
[3] *Institute of Cognitive Neuroscience, University College London, London, UK*

Edited by:
Leonhard Schilbach, University Hospital Cologne, Germany

Reviewed by:
Jean Decety, University of Chicago, USA
Chris Frith, Wellcome Trust Centre for Neuroimaging at University College London, UK

****Correspondence:***
Patricia L. Lockwood, Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London WC1H 0AP, UK
e-mail: p.lockwood@ucl.ac.uk

Individuals with psychopathy or autism spectrum disorder (ASD) can behave in ways that suggest lack of empathy towards others. However, many different cognitive and affective processes may lead to unempathic behavior and the social processing profiles of individuals with high psychopathic vs. ASD traits are likely different. Whilst psychopathy appears characterized by problems with resonating with others' emotions, ASD appears characterized by problems with cognitive perspective-taking. In addition, alexithymia has previously been associated with both disorders, but the contribution of alexithymia needs further exploration. In a community sample (N = 110) we show for the first time that although affective resonance and cognitive perspective-taking are related, high psychopathic traits relate to problems with resonating with others' emotions, but not cognitive perspective taking. Conversely, high ASD traits relate to problems with cognitive perspective-taking but not resonating with others' emotions. Alexithymia was associated with problems with affective resonance independently of psychopathic traits, suggesting that different component processes (reduced tendency to feel what others feel and reduced ability to identify and describe feelings) comprise affective resonance. Alexithymia was not associated with the reduced cognitive perspective-taking in high ASD traits. Our data suggest that (1) elevated psychopathic and ASD traits are characterized by difficulties in different social information processing domains and (2) reduced affective resonance in individuals with elevated psychopathic traits and the reduced cognitive perspective taking in individuals with elevated ASD traits are not explained by co-occurring alexithymia. (3) Alexithymia is independently associated with reduced affective resonance. Consequently, our data point to different component processes within the construct of empathy that are suggestive of partially separable cognitive and neural systems.

Keywords: psychopathy, autism spectrum disorder, alexithymia, empathy, affective resonance, cognitive perspective-taking

INTRODUCTION

Empathy is the capacity to understand or resonate with the affective experiences of others (Singer and Lamm, 2009). Two important processes that contribute to empathy are (i) being aware of, and resonating with, the feelings of another individual such that the awareness of their emotion drives the same state in oneself (henceforth affective resonance) and (ii) identifying and understanding what another individual is thinking/feeling without a necessary affective response (henceforth cognitive perspective-taking). These processes may differentially characterize psychopathy and autism spectrum disorders (ASDs). Although individuals with either disorder can behave in ways that suggest lack of empathy towards others' (Blair, 2005; Jones et al., 2010) this may be the result of problems in different social information processing domains.

Psychopathy is a disorder characterized by a lack of empathy, shallow affect, and manipulation of others for own gain (Hare, 2003). Difficulties with affective resonance are often apparent. For example, individuals with psychopathy show reduced physiological response to others' distress (Blair et al., 1997). Adults with psychopathy and children with psychopathic traits display atypical neural responses to others' pain (Decety et al., 2013; Lockwood et al., 2013; Marsh et al., 2013). In community samples, high levels of psychopathic traits are related to weaker affective responses to fearful faces and happy stories (Seara-Cardoso et al., 2012, 2013). Taken together, these findings indicate clear difficulties in resonating with others' emotions in both clinical samples with psychopathy and in community individuals with high levels of psychopathic traits. In contrast, one of the defining features of psychopathy is the ability to successfully manipulate others (Hare, 2003). Thus it might be expected that psychopathy would be associated with typical cognitive perspective-taking. Several studies report no cognitive perspective-taking impairments (Blair et al., 1996; Richell et al., 2003; Dolan and Fullam,

2004; Anastassiou-Hadjicharalambous and Warden, 2008) and even superior ability (Hansen et al., 2008) in individuals with psychopathy or high psychopathic traits. However, others have reported problems with tasks related to cognitive perspective-taking in both incarcerated psychopaths (Brook and Kosson, 2013) and healthy samples with high psychopathic traits (Ali and Chamorro-Premuzic, 2010). One possibility for these mixed findings is that different paradigms vary in their level of affective content, with some purported cognitive perspective-taking measures requiring identification of other people's feelings, rather than just their thoughts. It could be that negative associations between psychopathic traits and cognitive perspective-taking are driven by problems related to basic affective processing, rather than difficulties in cognitive perspective-taking *per se*. In fact, all studies that have reported that psychopathy/psychopathic traits are associated with poorer cognitive perspective-taking have utilized measures with affective content (e.g., Ali and Chamorro-Premuzic, 2010; Brook and Kosson, 2013) and therefore do not necessarily provide evidence for cognitive perspective-taking impairments in psychopathy.

Autism spectrum disorders are characterized by problems with social interaction, communication, and repetitive behaviors. ASD are also associated with atypical empathic processing (e.g., Baron-Cohen and Wheelwright, 2004). Several decades of research indicates that individuals with ASD have difficulties with cognitive perspective-taking (see Hill and Frith, 2003). The findings from studies assessing processes related to affective resonance in ASD are less consistent. There is evidence of absent sensori-motor resonance when viewing others' pain in individuals with ASD (Minio-Paluello et al., 2009). However, other studies have shown typical sensori-motor resonance when viewing others in pain (Fan et al., 2013) and appropriate physiological responses to others distress (Blair, 1999) in individuals with ASD. When cognitive perspective-taking and empathic concern, a process related to affective resonance, have been compared in individuals with ASD, impairments in cognitive perspective-taking but not empathic concern were found (Dziobek et al., 2008). Some theorists have argued that affective resonance is actually heightened in individuals with ASD (Smith, 2009) and reports of greater empathic facial affect in children with ASD compared to controls supports this (Capps et al., 1993).

A further consideration is the high comorbidity of ASD with alexithymia. Alexithymia is a sub-clinical condition defined by an inability to identify and describe feelings in the self. Preliminary behavioral and neuroimaging research suggests that affective and empathy impairments in ASD may be a function of interoceptive difficulties related to alexithymia rather than ASD *per se* (Silani et al., 2008; Bird et al., 2010) and that after accounting for alexithymia there is no difference in empathy between individuals with ASD and controls (Bird and Cook, 2013). However, one recent fMRI study found no significant moderating effects of alexithymia in an empathy for pain task in individuals with ASD (Fan et al., 2013). Nevertheless, the variance in alexithymia scores was very limited (SD 3.8 in Fan et al., 2013 vs. 11.8 in Bird et al., 2010), which may explain why no effect of alexithymia was observed. Less is known about the possible contribution of alexithymia to empathy impairments seen in individuals with psychopathy. Although the co-occurrence rates of alexithymia and psychopathy are lower than for ASD (Louth et al., 1998), the two disorders do share some common attributes (Lander et al., 2012).

To date, only two studies have directly compared the profile of affective and cognitive processing related to psychopathy and ASD, and these have both been in children. Children with conduct disorder and psychopathic traits showed less affective resonance with others' emotions but did not have problems with cognitive perspective-taking; conversely, children with ASD showed reduced cognitive perspective-taking but did not have problems with affective resonance (Jones et al., 2010; Schwenck et al., 2012). However, no studies have directly contrasted psychopathic and ASD traits and processes related to affective resonance and cognitive perspective-taking in adults. Moreover, no studies have investigated the contribution of alexithymia to ASD and psychopathic traits in tandem. Psychopathic, ASD and alexithymic traits are present in varying degrees in the general population (Bagby et al., 1994; Baron-Cohen et al., 2001; Hare and Neumann, 2008). Indeed, taxometric studies indicate that psychopathy should be viewed as a dimensional construct that is an extreme variant of normal personality and not a distinct category of behavior (see Hare and Neumann, 2008 for review). Similarly, behavioral genetic studies indicate a similar etiology of autistic traits in the general population as well as in clinical groups (Robinson et al., 2011), thus providing an empirical basis for studying variants in traits associated with these disorders in the general population. Finally, investigating associations between these traits and potential differences in social information processing is one way to dissect the component processes that may contribute to empathy.

Consequently, the present study investigated (i) whether psychopathic and ASD traits were differentially related to performance on affective resonance and cognitive perspective-taking tasks and (ii) whether alexithymia contributes to task performance. We predicted that psychopathic traits would be negatively associated with performance on the affective resonance task but not the cognitive perspective-taking task and that ASD traits would be negatively associated with performance on the cognitive perspective-taking task but not the affective resonance task. Alexithymia has previously been demonstrated to predict empathy deficits while recent neuroimaging results suggest cognitive perspective-taking is unlikely to be affected (Bernhardt et al., 2013). Therefore, we predicted that alexithymia would make a contribution to performance on the affective resonance task, but be unrelated to performance on the cognitive perspective-taking task. We also explored whether the proposed association with alexithymia would reflect variance common to alexithymia and psychopathic traits, or variance unique to alexithymia.

MATERIALS AND METHODS

PARTICIPANTS

One hundred and ten healthy adults (50% M; 50% F) aged 18–33 ($M = 21.9$, SD $= 3.7$) with estimated IQ between 87 and 129 ($M = 116.8$, SD $= 8.4$) took part. Participants were recruited through university participant databases and the community. All participants provided written informed consent and the study had institutional ethics approval.

PROCEDURE

Participants completed two tasks to assess affective resonance and cognitive perspective-taking as part of a larger battery of tasks. All tasks were presented in a randomized order followed by the questionnaires.

EXPERIMENTAL TASKS

Theory of mind animations task (cognitive perspective-taking task)

This task assessed participants' ability to understand others' complex mental states (e.g., tricking, coaxing) and has been previously used to examine ToM abilities in children with autism (Abell et al., 2000) and healthy participants (Castelli et al., 2002). We selected four "ToM" and four "random" animations from Abell et al. (2000). Each animation featured two characters; a big red and small blue triangle either interacting with one another (ToM animations) or moving randomly (random animations). Participants were asked to watch each animation carefully and to describe what was happening whilst their verbal responses were recorded. Two people transcribed the verbal descriptions that were coded in terms of intentionality and appropriateness. The intentionality scale ranged from 0 (no appreciation of another agent, nor actions or mental states) to 5 (the agent acts with the goal of affecting or manipulating the other agent's mental states). The appropriateness scale ranged from 0 to 3. One researcher rated all transcriptions and a second researcher rated a random sample of 56. Intra-class correlations (ICC) between raters for intentionality (ICC, single measures = 0.682) and appropriateness (ICC single measures = 0.760) were good. The ratings of intentionality and appropriateness were converted to *z*-scores and a composite score was created.

Self-assessment manikin faces task (Affective resonance task)

This task assessed participants' affective empathic response to emotional faces using the SAM rating scale (Seara-Cardoso et al., 2012). Participants were required to rate their own emotional response to the affective state of another on a nine-point manikin (changing from smiling to a sad face with a neutral expression in the middle) whilst viewing images depicting a person showing either a sad, fearful, angry, happy, or neutral expression. The order of images was randomized for each participant. Ratings for sad, fear, and anger were reverse scored so that the higher scores reflected ratings of greater distress, and thus greater affective resonance, when viewing others' negative emotions. These variables were then converted to *z*-scores and a composite score was created along with happy ratings.

QUESTIONNAIRES

Self-Report Psychopathy Scale–Short Form (SRP-4-SF, Paulhus et al., in press)

Psychopathic traits were assessed with the SRP-4-SF, a 29-item scale designed to measure psychopathic attributes in non-institutionalized samples. The SRP has been shown to have good construct validity and internal consistency (Cronbach's alpha 0.89 in the present study) and is strongly correlated with the PCL-R; the clinical measure of psychopathy (Lilienfeld and Fowler, 2006; Paulhus et al., in press). Questions were rated on a five-point scale from "Disagree Strongly" to "Agree Strongly" and included items such as "Most people are wimps" and "I love violent sports and movies."

The Autism Spectrum Quotient (AQ, Baron-Cohen et al., 2001)

Autism spectrum disorder traits were assessed with the AQ, a 50-item scale designed to assess ASD traits in both clinical and community samples. The AQ has good construct validity and internal consistency (Cronbach's alpha 0.83 in the present study). Questions were rated on a four-point scale from "Definitely Disagree" to "Definitely Agree" and included items such as "I enjoy meeting new people" and "I would rather go to a library than a party."

Toronto Alexithymia scale (TAS, Bagby et al., 1994)

Alexithymic traits were assessed with the TAS, a 20-item scale designed to measure subclinical alexithymic traits. Questions were rated on a five-point scale from "I Strongly Disagree" to "I Strongly Agree" and included items such as "I am often confused about what emotion I am feeling" and "I am often puzzled by sensations in my body." The TAS has good construct validity and internal consistency (Cronbach's alpha 0.82 in the present study).

RESULTS

Performance on the affective resonance and cognitive perspective-taking tasks was positively correlated ($r = 0.40$, $p < 0.001$). All questionnaire measures were also positively correlated with one another (see **Table 1**). First, bivariate correlations were examined to assess whether psychopathic and ASD traits were differentially related to affective resonance and cognitive perspective-taking. As predicted psychopathic traits showed a statistically significant negative correlation with performance on the affective resonance task ($r = -0.258$, $p = 0.007$) whilst ASD traits did not ($r = -0.102$, $p = 0.291$). Conversely, ASD traits showed a statistically significant negative correlation with performance on the cognitive perspective-taking task ($r = -0.209$, $p = 0.028$) whilst psychopathic traits did not ($r = -0.046$, $p = 0.634$).

We conducted hierarchical multiple regression analyses to investigate whether psychopathic and ASD traits were uniquely and differentially related to affective resonance and cognitive perspective-taking, and to examine whether individual differences in alexithymia and/or IQ might explain any associations

Table 1 | Correlations between questionnaire measures of psychopathic, autism spectrum disorder, and alexithymic traits and task performance.

	SRP	AQ	TAS	AR
AQ	0.244*			
TAS	0.252*	0.370**		
AR	−0.258**	−0.102	−0.245*	
CPT	−0.046	−0.209*	−0.120	0.399**

*SRP, Self-Report Psychopathy Scale; TAS, Toronto Alexithymia Scale; AQ, Autism Spectrum Quotient; AE, affective resonance task; CPT, cognitive perspective-taking task. *p < 0.05, **p < 0.01.*

Table 2 | Hierarchical multiple regression between questionnaire measures of psychopathic, autism spectrum disorder, and alexithymic traits and task performance.

	Affective resonance task				Cognitive perspective-taking task		
	Beta	***t***	***P***		**Beta**	***t***	***P***
STEP 1							
SRP	−0.258	−2.772	0.007*	AQ	−0.209	−2.224	0.028*
STEP 2							
SRP	−0.248	−2.574	0.011*	AQ	−.211	−2.16	0.033*
AQ	−0.041	−0.428	0.669	SRP	0.005	0.056	0.956
STEP 3							
SRP	−0.213	−2.209	0.029*	AQ	−0.193	−1.868	0.065^
AQ	0.025	0.245	0.807	SRP	0.014	0.144	0.885
TAS	−0.201	−1.991	0.049*	TAS	−0.052	−0.501	0.618
STEP 4							
SRP	−0.218	−2.227	0.028*	AQ	−0.196	−1.895	0.061^
AQ	0.024	0.236	0.814	SRP	−0.000	0.000	1.000
TAS	−0.200	−1.977	0.051^	TAS	−0.050	−0.483	0.630
IQ	0.033	0.353	0.725	IQ	0.106	1.113	0.268

^$p < 0.10$, *$p < 0.05$.
SRP, Self-Report Psychopathy Scale; TAS, Toronto Alexithymia Scale; AQ, Autism Spectrum Quotient; Full IQ calculated from Weschler Intelligence Test of Adult Reading.

(see **Table 2**). Two models were run. For the model predicting performance on the affective resonance task, psychopathic traits were entered at the first stage. Psychopathic traits significantly predicted reduced affective resonance ($p = 0.007$). At the second stage ASD traits were entered. Psychopathic traits were uniquely negatively associated with affective resonance ($t = -2.57$, $p = 0.011$) whilst ASD traits were not ($t = -0.43$, $p = 0.669$). The R^2 change was not significant (F change $= 0.18$, $p = 0.669$) indicating that ASD traits did not significantly explain more variance in the model. At the third stage, alexithymia scores were entered. Controlling for alexithymia did not change the pattern of results, but there was a unique negative association between alexithymia and affective resonance ($t = -1.99$, $p = 0.049$), and the R^2 change was significant ($F = 3.96$, $p = 0.049$). At the fourth stage IQ scores were entered. Controlling for IQ did not change the pattern of results, nor was IQ a significant predictor of affective resonance ($p = 0.73$). The same regression sequence was then used for cognitive perspective-taking, but with ASD traits at the first stage and psychopathic traits at the second. ASD traits were significantly negatively associated with cognitive perspective-taking ($t = -2.22$, $p = 0.028$). At the second stage psychopathic traits were entered. ASD traits were uniquely negatively associated with reduced cognitive perspective taking ($t = -2.16$, $p = 0.033$) whilst psychopathic traits were not ($t = 0.06$, $p = 0.956$). The R^2 change was not significant (F change $= 0.00$, $p = 0.956$) indicating that psychopathic traits did not explain significantly more variance in the model. Taking into account alexithymia and IQ did not change the pattern of results, nor did either of these variables predict cognitive perspective-taking. No further R^2 changes were significant (all F's < 1.24, all ps > 0.26).

DISCUSSION

The current study compared associations between psychopathic or ASD traits and tasks assessing affective resonance or cognitive perspective-taking. We demonstrated unique associations between psychopathic traits and reduced affective resonance but not cognitive perspective-taking, and unique associations between ASD traits and reduced cognitive perspective-taking but not affective resonance. Alexithymic traits did not explain observed associations between task performance and psychopathic or ASD traits but rather contributed to performance on the affective resonance task independently of psychopathic traits. This is the first study in healthy adults to show a differential relationship between these variables. Thus, it extends previous findings that have reported contrasting profiles of empathy impairments between children with psychopathic tendencies or ASD (Jones et al., 2010; Schwenck et al., 2012). Our results also suggest that although affective resonance and cognitive perspective-taking measures share some variance, they can capture dissociable processes.

Psychopathy is thought to be characterized by problems with affective resonance but not cognitive perspective-taking. We used measures that were designed to specifically probe affective resonance and cognitive perspective-taking, without there being cognitive perspective-taking demands on the affective resonance task or vice versa. Our results therefore extend and clarify the findings of previous studies reporting reduced affective resonance in individuals high in psychopathic traits (Seara-Cardoso et al., 2012, 2013) by indicating a reduction in affective resonance in the absence of a reduction in cognitive perspective-taking. These data also highlight how high psychopathic

traits are not related to atypical cognitive perspective-taking processing when a task without an affective component is used.

Autism spectrum disorders have been consistently linked to problems with cognitive perspective-taking (Hill and Frith, 2003). Interestingly, we found that elevated ASD traits in the general population were also associated with atypical cognitive perspective-taking. In contrast, findings of tasks related to affective resonance processing in autism are mixed, with reduced (Minio-Paluello et al., 2009), intact (Blair, 1999; Dziobek et al., 2008; Bird et al., 2010; Fan et al., 2013), and elevated (Capps et al., 1993) levels of affective processing being reported. Our findings suggest that ASD traits are not associated with either a reduced or an enhanced ability to resonate with the emotions of another, despite the fact that high levels of ASD traits are related to difficulties with understanding others' minds. It would be useful for future studies to assess multiple forms of processing related to affective resonance, as the paradigms used in some studies that reported intact affective resonance investigated empathic concern, rather than affective resonance. Examining both of these processes in tandem may help to shed further light on the profile of empathic processing in ASD. Moreover, it would also be interesting to further examine the exact cognitive perspective-taking mechanisms that may be disrupted in relation to ASD/high ASD traits. It could be that some disrupted components of cognitive perspective-taking relate to bottom–up processes such as detection of biological movement, whereas others might relate to top–down processes such as the influence of situational cues.

Both psychopathy and ASD have previously been associated with elevated levels of alexithymia (Louth et al., 1998; Lander et al., 2012; Bird and Cook, 2013), and we also observed modest correlations between psychopathic and ASD traits with alexithymia in the present study. Nevertheless, controlling for alexithymic traits did not change the reported associations between psychopathic traits and reduced affective resonance or ASD traits and reduced cognitive perspective-taking. In other words, the reduced ability to identify and describe feelings in the self did not account for the relationship between psychopathic traits and affective resonance or ASD traits and cognitive perspective-taking. The finding that alexithymia did not explain the reduced cognitive perspective-taking abilities characteristic of ASD traits is of particular interest given recent evidence and theory suggesting that alexithymia does account for affective processing deficits related to autism, when they are observed (Bird and Cook, 2013). Our data extend this account by showing that alexithymia does not appear to explain reduced cognitive perspective-taking related to high ASD traits.

We also found that alexithymic traits were negatively associated with a reduction in affective resonance independently of psychopathic traits. This suggests that reductions in affective resonance can be affected both by reduced ability to identify and describe feelings (a characteristic of alexithymia) and a reduced tendency to feel what others feel (a characteristic of psychopathy). The result of independence between psychopathic and alexithymic traits in predicting performance on affective resonance also points to potential component processes within the construct of affective resonance. Future studies could help to determine the mechanisms underlying reduced affective resonance in psychopathy and alexithymia.

A few limitations to the present study should be highlighted. In everyday life empathic responses to others occur in the context of reciprocal social interactions, the present tasks did not present such scenarios in the interest of isolating affective resonance and cognitive perspective-taking demands. Although we chose paradigms to specifically examine two process that contribute to the experience of empathy, these are not exhaustive and further research would benefit from examining a larger collection of tasks that tap a multitude of processes related to empathy. It will also be of interest to determine whether the processing atypicalities associated with psychopathic, ASD, and alexithymia traits explain real life observations of unempathic behavior, as rated by others or observed in an experimental setting. Finally, replication of these results with clinical populations would be informative.

Overall, our findings clarify and extend previous studies examining the profiles of empathy deficits related to psychopathy, ASD, and alexithymia. We show for the first time that in subclinical samples elevated psychopathic traits are related to reduced affective resonance but not cognitive perspective-taking whilst elevated levels of ASD traits are related to reduced cognitive perspective-taking but not affective resonance. Consequently, our data point to different social information processes within the construct of empathy that are suggestive of partially separable cognitive and neural systems.

ACKNOWLEDGMENTS

This work was supported by a Doctoral Training Account studentship from the Medical Research Council awarded to Patricia L. Lockwood. Essi Viding is a Royal Society Wolfson Research Merit Award holder. Geoffrey Bird was supported by a grant from the Baily Thomas Charitable Fund (3089/1). Geoffrey Bird completed the paper while a Senior Fellow at the Netherlands Institute for Advanced Study in the Humanities and Social Sciences. We would like to thank Prof. Craig Neumann for his help and advice with regard to the Self-Report Psychopathy Scale.

AUTHORS CONTRIBUTION

Patricia L. Lockwood, Geoffrey Bird, and Essi Viding designed research, Patricia L. Lockwood and Madeleine Bridge collected data. Patricia L. Lockwood analyzed data, Patricia L. Lockwood, Geoffrey Bird, and Essi Viding wrote paper.

REFERENCES

Abell, F., Happe, F., and Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cogn. Dev.* 15, 1–16. doi: 10.1016/S0885-2014(00)00014-9

Ali, F., and Chamorro-Premuzic, T. (2010). Investigating Theory of Mind deficits in nonclinical psychopathy and Machiavellianism. *Pers. Individ. Dif.* 49, 169–174. doi: 10.1016/j.paid.2010.03.027

Anastassiou-Hadjicharalambous, X., and Warden, D. (2008). Cognitive and affective perspective-taking in conduct-disordered children high and low on callous-unemotional traits. *Child Adolesc. Psychiatry Ment. Health* 2, 16. doi: 10.1186/1753-2000-2-16

Bagby, R. M., Parker, J. D. A., and Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *J. Psychosom. Res.* 38, 23–32. doi: 10.1016/0022-3999(94)90005-1

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/A:1005653411471

Bernhardt, B. C., Valk, S. L., Silani, G., Bird, G., Frith, U., and Singer, T. (2013). Selective disruption of sociocognitive structural brain networks in autism and alexithymia. *Cereb Cortex.* doi:10.1093/cercor/bht182 [Epub ahead of print].

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., and Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain* 133, 1515–1525. doi: 10.1093/brain/awq060

Bird, G., and Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Transl. Psychiatry* 3, e285. doi: 10.1038/tp.2013.61

Blair, R, J. R., Jones, L., Clark, F., and Smith, M. (1997). The psychopathic individual: a lack of responsiveness to distress cues? *Psychophysiology* 34, 192–198. doi: 10.1111/j.1469-8986.1997.tb02131.x

Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., et al. (1996). Theory of mind in the psychopath. *J. Forens. Psychiatry* 7, 15–25. doi: 10.1080/09585189608409914

Blair, R. J. R. (1999). Psychophysiological responsiveness to the distress of others in children with autism. *Pers. Individ. Dif.* 26, 477–485. doi: 10.1016/S0191-8869(98)00154-8

Blair, R. J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718. doi: 10.1016/j.concog.2005.06.004

Brook, M., and Kosson, D. S. (2013). Impaired cognitive empathy in criminal psychopathy: evidence from a laboratory measure of empathic accuracy. *J. Abnorm. Psychol.* 122, 156. doi: 10.1037/a0030261

Capps, L., Kasari, C., Yirmiya, N., and Sigman, M. (1993). Parental perception of emotional expressiveness in children with autism. *J. Consult. Clin. Psychol.* 61, 475. doi: 10.1037/0022-006X.61.3.475

Castelli, F., Frith, C., Happé, F., and Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125, 1839–1849. doi: 10.1093/brain/awf189

Decety, J., Skelly, L., and Kiehl, K. (2013). Brain response to empathy-eliciting scenarios involving pain in incarcerated individuals with psychopathy. *JAMA Psychiatry* 70, 638–645. doi: 10.1001/jamapsychiatry.2013.27

Dolan, M., and Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychol. Med.* 34, 1093–1102. doi: 10.1017/S0033291704002028

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., et al. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *J. Autism Dev. Disord.* 38, 464–473. doi: 10.1007/s10803-007-0486-x

Fan, Y. T., Chen, C., Chen, S. C., Decety, J., and Cheng, Y. (2013). Empathic arousal and social understanding in individuals with autism: evidence from fMRI and ERP measurements. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst101 [Epub ahead of print].

Hansen, A. L., Johnsen, B. H., Hart, S., Waage, L., and Thayer, J. F. (2008). Brief communication: psychopathy and recognition of facial expressions of emotion. *J. Pers. Disord.* 22, 639–644. doi: 10.1521/pedi.2008.22.6.639

Hare, R. (2003). *Hare Psychopathy Checklist – Revised.* Toronto, ON: Multi-Health Systems.

Hare, R. D., and Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annu. Rev. Clin. Psychol.* 4, 217–246. doi: 10.1146/annurev.clinpsy.3.022806.091452

Hill, E. L., and Frith, U. (2003). Understanding autism: insights from mind and brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 281–289. doi: 10.1098/rstb.2002.1209

Jones, A. P., Happé, F. G., Gilbert, F., Burnett, S., and Viding, E. (2010). Feeling, caring, knowing: different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *J. Child Psychol. Psychiatry* 51, 1188–1197. doi: 10.1111/j.1469-7610.2010.02280.x

Lander, G. C., Lutz-Zois, C. J., Rye, M. S., and Goodnight, J. A. (2012). The differential association between alexithymia and primary versus secondary psychopathy. *Pers. Individ. Dif.* 52, 45–50. doi: 10.1016/j.paid.2011.08.027

Lilienfeld, S. O., and Fowler, K. A. (2006). "The self-report assessment of psychopathy: problems, pitfalls, and promises," in *Handbook of psychopathy*, ed. C. J. Patrick (New York: Guilford Press), 107–132.

Lockwood, P. L., Sebastian, C. L., McCrory, E. J., Hyde, Z. H., Gu, X., De Brito, S. A., et al. (2013). Association of callous traits with reduced neural response to others' pain in children with conduct problems. *Curr. Biol.* 23, 901–905. doi: 10.1016/j.cub.2013.04.018

Louth, S. M., Hare, R. D., and Linden, W. (1998). Psychopathy and alexithymia in female offenders. *Can. J. Behav. Sci.* 30, 91. doi: 10.1037/h0085809

Marsh, A. A., Finger, E. C., Fowler, K. A., Adalio, C. J., Jurkowitz, I. T., Schechter, J. C., et al. (2013). Empathic responsiveness in amygdala and anterior cingulate cortex in youths with psychopathic traits. *J. Child Psychol. Psychiatry* 54, 900–910. doi: 10.1111/jcpp.12063

Minio-Paluello, I., Baron-Cohen, S,. Avenanti, A., Walsh, V., and Aglioti, S. M. (2009). Absence of embodied empathy during pain observation in Asperger syndrome. *Biol. Psychiatry* 65, 55–62. doi: 10.1016/j.biopsych.2008.08.006

Paulhus, D. L., Newman, C., and Hare, R. (in press). *Manual for the Self-Report Psychopathy Scale.* Toronto, ON: Multi-Health systems.

Richell, R. A., Mitchell, D. G. V., Newman, C., Leonard, A., Baron Cohen, S., and Blair, R. J. R. (2003). Theory of mind and psychopathy: can psychopathic individuals read the 'language of the eyes'? *Neuropsychologia* 41, 523–526. doi: 10.1016/S0028-3932(02)00175-6

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happe, F., et al. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Arch. Gen. Psychiatr.* 68, 1113–1121. doi: 10.1001/archgenpsychiatry.2011.119

Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., et al. (2012). Empathy in children with autism and conduct disorder: group-specific profiles and developmental aspects. *J. Child Psychol. Psychiatry,* 53, 651–659. doi: 10.1111/j.1469-7610.2011.02499.x

Seara-Cardoso, A., Craig, N., Roiser, J., McCrory, E., and Viding, E. (2012). Investigating associations between empathy, morality and psychopathic personality traits in the general population. *Pers. Individ. Dif.* 52, 67–71. doi: 10.1016/j.paid.2011.08.029

Seara-Cardoso, A., Dolberg, H., Neumann, C., Roiser, J., and Viding. E. (2013). Empathy, morality and psychopathic traits in women. *Pers. Individ. Dif.* 55, 328–333. doi: 10.1016/j.paid.2013.03.011

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., and Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Soc. Neurosci.* 3, 97–112. doi: 10.1080/17470910701577020

Singer, T., and Lamm, C. (2009). The social neuroscience of empathy. *Ann. N. Y. Acad. Sci.* 1156, 81–96. doi: 10.1111/j.1749-6632.2009.04418.x

Smith, A. (2009). The empathy imbalance hypothesis of autism: a theoretical approach to cognitive and emotional empathy in autistic development. *Psychol. Rec.* 3, 9.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 September 2013; accepted: 22 October 2013; published online: 13 November 2013.

*Citation: Lockwood PL, Bird G, Bridge M and Viding E (2013) Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. Front. Hum. Neurosci. **7**:760. doi: 10.3389/fnhum.2013.00760*

This article was submitted to the journal Frontiers in Human Neuroscience.

OPEN ACCESS Freely available online

PLOS ONE

Emotion Regulation Moderates the Association between Empathy and Prosocial Behavior

Patricia L. Lockwood*, Ana Seara-Cardoso, Essi Viding

Division of Psychology and Language Sciences, University College London, London, United Kingdom

Abstract

Theory and evidence suggest that empathy is an important motivating factor for prosocial behaviour and that emotion regulation, i.e. the capacity to exert control over an emotional response, may moderate the degree to which empathy is associated with prosocial behaviour. However, studies to date have not simultaneously explored the associations between different empathic processes and prosocial behaviour, nor whether different types of emotion regulation strategies (e.g. cognitive reappraisal and expressive suppression) moderate associations between empathy and prosocial behaviour. One hundred–and-ten healthy adults completed questionnaire measures of empathy, emotion regulation and prosocial tendencies. In this sample, both affective and cognitive empathy predicted self-reported prosocial tendencies. In addition, cognitive reappraisal moderated the association between affective empathy and prosocial tendencies. Specifically, there was a significant positive association between empathy and prosocial tendencies for individuals with a low or average tendency to reappraise but not for those with a high tendency to reappraise. Our findings suggest that, in general, empathy is positively associated with prosocial behaviour. However, this association is not significant for individuals with a high tendency for cognitive reappraisal.

Citation: Lockwood PL, Seara-Cardoso A, Viding E (2014) Emotion Regulation Moderates the Association between Empathy and Prosocial Behavior. PLoS ONE 9(5): e96555. doi:10.1371/journal.pone.0096555

Editor: Matjaz Perc, University of Maribor, Slovenia

Received February 25, 2014; **Accepted** April 9, 2014; **Published** May 8, 2014

Funding: This work was supported by a Doctoral Training Account studentship from the Medical Research Council awarded to Patricia L. Lockwood. Essi Viding is a Royal Society Wolfson Research Merit Award holder. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: p.lockwood@ucl.ac.uk

Introduction

Humans have a remarkable capacity to engage in prosocial behaviours, i.e. social behaviour intended to benefit another, with genetically unrelated individuals [1]. However, the processes that influence how and when prosocial behaviours occur remain poorly understood. Theory and evidence suggest that empathy, i.e. the capacity to understand and/or resonate with the affective experiences of others [2], is one of the key motivating factors for prosocial behaviour [3–5].

A number of processes are thought to contribute to the experience of empathy. These include 'affective' empathic processes, such as being aware of and resonating with the feelings of another individual, as well as 'cognitive' empathic processes, such as identifying and understanding what another individual is thinking or feeling without a necessary affective response [1]. There is evidence that processes related to affective and cognitive empathy are positively associated with prosocial behaviour (for a review see [3]). The majority of these studies have used the interpersonal reactivity index (IRI, [6]), which measures dispositional empathic concern/sympathy, or cardiovascular and electrodermal indices, such as heart rate deceleration and facial electromyographic (EMG), as proxy measures of affective empathy. For example, heart rate deceleration (which is thought to index vicariously induced sadness or sympathy, e.g [7]) and increased indicators of facial sadness when watching needy others are associated with increased willingness to help [8]. Dispositional empathic concern, as measured by the IRI, has also been linked to higher levels of self-reported charitable giving [9] and greater self-reported concern for the welfare of others [10]. In terms of associations between cognitive components of empathy and prosocial behaviour, studies have focused on correlating the perspective-taking subscale of the IRI to self-reported prosocial behaviour and have found that trait perspective taking is positively associated with frequency of volunteering [11] and self-reported prosocial tendencies [12]. It should be noted, however, that the empathic concern and perspective taking scales of the IRI tap constructs that, although related, are different from the current conceptualisation of 'affective' and 'cognitive empathy' [2]. Nonetheless, together, these studies broadly suggest that affective and cognitive empathic processes may motivate prosocial behaviour.

Whilst it is often assumed that an empathic response to another's distress will motivate prosocial behaviour, Eisenberg [13] points out that association between the two constructs are often modest and sometimes weak. A possible reason for these modest associations is the influence of moderating variables [13]. It has been suggested that emotion regulation, i.e. the capacity to modulate or exert control over an emotional response, might be one such moderator variable [14], [15]. Eisenberg and Fabes [14] propose a model whereby individual differences in both the emotional intensity and regulation capacities are related to an individual's level of prosocial responding. Specifically, they suggest that the perception of distress in another leads to emotional arousal, but emotion regulation i.e. and how this arousal is evaluated by the observer, will influence the subsequent goal

directed behaviour, either to improve their own situation or help the others' situation [14]. The degree of emotion regulation during a state of emotional arousal (over-, optimal-, or under-regulation) is also proposed to relate to the likelihood of prosocial behaviour. For example, individuals who are able to optimally regulate their arousal, so that they do not experience undue distress in the face of another person's emotions and thus do not become self-focused, are proposed to behave prosocially [14]. In contrast, individuals who are over- regulated are proposed to exhibit proactive withdrawal, which inhibits prosocial behaviour. Finally, those who are under-regulated are proposed to be prone to aggression and thus more likely to exhibit antisocial rather than prosocial behaviour in an emotionally arousing situation [14].

The model outlined by Eisenberg and Fabes [14] discusses the degree of emotion regulation (over-, optimal-, or under-regulation) as important for linking empathy to prosocial behaviour. However, it is also likely that the type of emotion regulation strategy used will be important. Both cognitive reappraisal and expressive suppression represent emotion regulation strategies [16–18]. Cognitive reappraisal involves reinterpreting an emotional response so that the intensity of its emotional impact is modified [19]. For example, re-framing a distressing situation as a situation where someone will benefit from support, as opposed to a situation where someone is emotionally labile and potentially unpleasant. Consequently, cognitive reappraisal will enable a person to focus on strategies to provide constructive helping behaviours, rather than the aversive qualities of the situation. Cognitive reappraisal is thought to be a successful emotion regulation strategy, decreasing negative affect and resulting in an attenuation of blood pressure [20], [21]. In contrast, expressive suppression involves actively inhibiting on-going emotion-expressive behaviour [17], [18], [22]. For example, managing an emotional response to an aversive situation in an effortful manner such that cognitive resources are consumed. Expressive suppression is thought to be a suboptimal strategy because it creates a conflict between heightened emotional arousal and overt expression of the arousal [17], [18], [23]. These two types of emotion regulation strategies also appear to lead to different outcomes and consequences for interpersonal functioning [16], [24–26]. Whilst cognitive reappraisal is positively related to having closer relationships with friends, fewer depressive symptoms and greater life satisfaction, expressive suppression is associated with greater experience of negative emotions, disturbed interpersonal interactions, avoidance of close relationships and reports of less life satisfaction and optimism [16], [24–26].

Despite evidence linking empathy to prosocial behaviour (e.g. [8], [11]) and the proposal that individual differences in emotion regulation may moderate associations between empathy and prosocial behaviour [14], [15], this has not, to our knowledge, been directly examined. Moreover, how distinct emotion regulation strategies might moderate associations between empathy and prosocial behaviour has not been explored. The majority of studies suggesting empathy as a motivating factor for prosocial behaviour have investigated self-reported empathic concern (feeling 'for' another person, including compassion and sympathy, e.g. [9], [10]), rather than self-reported affective empathic responses (the ability to vicariously experience the emotional experience of others; or feeling 'as' another individual). While these two processes are no doubt closely related, there is a lack of empirical data regarding how feeling in a similar emotional state to another may motivate prosocial behaviour. In addition, self-reported cognitive empathic ability (i.e. the ability to position oneself 'in another person's shoes') might also relate to prosocial behaviour, but compared to the role of affective empathic processes motivating empathy this has received relatively little attention to date (c.f. [11], [12]).

On the basis of previous research and theory (e.g. [3], [10], [12]), we predicted that both dispositional cognitive and affective components of empathy would be associated with increased prosocial tendencies, but the amount of variance in prosocial behaviour explained by the two types of empathy may be unequal. We also tested interactions between the components of empathy (affective and cognitive) and types of emotion regulation strategy (cognitive reappraisal and expressive suppression) to examine whether individual differences in emotion regulation strategy moderate associations between empathy and prosocial behaviour.

Methods

Participants

One-hundred-and-ten healthy adults (50% males; 50% females) aged 18–33 (M = 21.9, SD = 3.7) were recruited through university participant databases (comprised of undergraduate and postgraduate students as well as non-student community members) and through online advertisement. Exclusion criteria included previous or current neurological or psychiatric disorder (as reported by the participants) and non-normal or non-corrected to normal vision. Participants were compensated at a rate of £8 per hour.

Ethics Statement

All participants provided written informed consent and the study was approved by the University College London Clinical, Educational and Health Psychology Research Ethics committee.

Procedure

Participants completed questionnaires to assess empathy, emotion regulation and prosocial tendencies as part of a larger battery of tasks and questionnaires.

Questionnaires

Questionnaire of Cognitive and Affective Empathy (QCAE; [27]). The QCAE, is a multidimensional empathy questionnaire devised to measure the ability to comprehend the emotions of another (cognitive empathy) as well as the ability to vicariously experience the emotional experience of others (affective empathy). In the development of the QCAE, two raters selected items from other well-validated and commonly used empathy measures (e.g. Empathy Quotient; [28], Hogan Empathy Scale; [29], the Empathy subscale of the Impulsiveness-Venturesomeness-Empathy Inventory; [30], and the IRI; [9]) if they were deemed to measure affective or cognitive empathy. Items from these scales deemed to measure other processes (e.g. sympathy) were not included. A Principal Component Analysis of the selected items revealed five components (or sub-scales), further organized in two dimensions assessing cognitive and affective empathy. The cognitive empathy dimension comprises subscales measuring perspective-taking (e.g. "I am good at predicting how someone will feel") and Online simulation (e.g. "Before criticizing somebody, I try to imagine how I would feel if I was in their place."). The affective subscales assess emotion contagion (e.g. "People I am with have a strong influence on my mood"); peripheral responsivity (e.g. "I usually stay emotionally detached when watching a film"); and proximal responsivity (e.g. "I often get emotionally involved with my friends' problems"). Items are rated on a 4-point scale from "strongly disagree" to "strongly agree". The QCAE has good validity and internal consistency [27]. In the present study Cronbach's alpha for cognitive empathy subscale .87; affective empathy subscale .88).

Emotion Regulation Questionnaire (ERQ; [19]). The ERQ is comprised of two dimensions that assess either reappraisal or suppression regulation strategies. The reappraisal dimension contains items such as "I control my emotions by changing the way I think about the situation I'm in" and the suppression dimension has items such as "I control my emotions by not expressing them". Items are rated on a 7-point scale from "Strongly disagree" to "Strongly agree". The ERQ has good construct validity and internal consistency ([19]; in the present study Cronbach's alpha for reappraisal subscale. 73; suppression subscale. 87).

Prosocial Tendencies Measure (PTM; [31]). The PTM is a 23-item self-report measure that assesses various prosocial tendencies such as compliant prosocial tendencies (e.g. "When people ask me to help them, I don't hesitate"), dire prosocial tendencies (e.g. "I tend to help people who hurt themselves badly") and emotional prosocial tendencies (e.g. "I tend to help others particularly when they are emotionally distressed"). Items are rated on a 5-point scale from "Does not describe me at all" to "Describes me greatly". The PTM has good construct validity and internal consistency ([31]; in the present study Cronbach's alpha .86).

Data Analyses

Bivariate correlations were corrected for multiple comparisons using Benjamini & Hochberg False Discovery Rate [32]. Corrected p-values are reported. Steiger's Z tests (two-tailed) were conducted to test if the different types of empathy (i.e. affective and cognitive empathy) and the different types of emotion regulation strategies (i.e. cognitive reappraisal and expressive suppression) presented differential correlation coefficients with prosocial tendencies.

Moderation analyses were then conducted to investigate whether the affective or cognitive empathy subscales interacted with either types of emotion regulation (reappraisal or suppression) to predict prosocial tendencies. All predictor variables were mean centred prior to analyses. Separate regression models using either the affective empathy subscale of the QCAE (QCAE-affective empathy) or the cognitive empathy subscale of the QCAE (QCAE-cognitive empathy) at the first stage; the reappraisal subscale of the ERQ (ERQ-reappraisal) or the suppression subscale of the ERQ (ERQ-suppression) at the second stage; the interaction term between these variables at the third stage were run. Consequently, four regression models were examined. Interaction effects were tested in SPSS using PROCESS [33]. Significant interactions were followed up by examining the conditional effect of empathy on prosocial tendencies at 1 standard deviation (SD) below the mean, at the mean, and 1 SD above the mean of emotion regulation.

Results

Bivariate correlations between questionnaire measures of empathy, emotion regulation and prosocial behaviour were examined (see Table 1 for a full list of correlations). QCAE-affective empathy and QCAE-cognitive empathy were both positively associated with prosocial tendencies ($r = .36$, $p < .001$ & $r = .43$, $p < .001$ respectively) and these correlations were not significantly different ($z = -.80$, $p > .05$). ERQ-reappraisal was not significantly correlated with prosocial tendencies ($r = .11$, $p = .30$). ERQ-suppression was significantly negatively correlated with prosocial tendencies ($r = -.27$, $p = .006$). These two correlations were significantly different ($Z = 2.69$, $p < .05$).

To examine whether the associations between affective and cognitive empathy and prosocial behaviour were explained by joint variance between the two components or whether they uniquely predicted prosocial tendencies we ran an additional multiple regression analysis. There were unique associations between each empathy component and prosocial tendencies (affective empathy, $t = 2.29$, $p = .024$; cognitive empathy, $t = 3.67$, $p < .001$).

For the first regression model we entered QCAE-affective empathy (first stage), ERQ-reappraisal (second stage), and their interaction term [QCAE-affective empathy×ERQ-reappraisal] (third stage) as predictors of prosocial tendencies. This analysis revealed a significant positive association between QCAE-affective empathy and prosocial tendencies ($t = 3.98$, $p < .001$) but not between reappraisal and prosocial tendencies ($t = .57$, $p = .570$). Interestingly, the interaction between QCAE-affective empathy and ERQ-reappraisal was significant ($t = -2.39$, $p = .019$). At 1 SD below the mean on ERQ-reappraisal there was a significant positive association between QCAE-affective empathy and prosocial tendencies ($t = 4.56$, $p < .001$). There was also a significant association at the mean ($t = 3.27$, $p = .002$). However at 1 SD above the mean on ERQ-reappraisal the association between QCAE-affective empathy and prosocial tendencies was non-significant ($t = 1.08$, $p = .282$) (see Figure 1). In other words, affective empathy was associated with prosocial behaviour for those with low and average levels of cognitive reappraisal (with the steepest slope for individuals with lowest level of cognitive appraisal), but those with high levels of cognitive reappraisal presented similar levels of prosocial behaviour regardless of level of affective empathy.

For the second regression model, QCAE-cognitive empathy, ERQ-reappraisal and their interaction term were entered as predictors of prosocial tendencies. This analysis showed a significant positive association between QCAE-cognitive empathy and prosocial tendencies ($t = 5.00$, $p < .001$) but not between reappraisal and prosocial tendencies ($t = -.39$, $p = .699$). The interaction between QCAE-cognitive empathy and ERQ-reappraisal was not significant ($t = -1.18$, $p = .243$). This pattern of findings suggests that QCAE-cognitive empathy was positively associated with prosocial tendencies regardless of level of reappraisal emotional regulation strategies.

We also examined the interaction between the two QCAE subscales and ERQ-suppression and their association with prosocial tendencies. These two regression models showed that both QCAE-AE and QCAE-CE were positively associated with prosocial tendencies ($t = 3.98$, $p < .001$ and $t = 5.00$, $p < .001$) but ERQ-suppression was not significantly associated with prosocial tendencies in either model ($t = -1.00$, $p = .32$ and $t = -1.36$, $p = .18$). Neither of the interactions between QCAE-affective empathy or QCAE-cognitive empathy and ERQ-suppression were significant (both $ps > .05$).

Discussion

The present study investigated associations between empathy and prosocial behaviour, and whether different types of emotion regulation strategy moderate associations between empathy and prosocial behaviour. We found that both affective and cognitive components of empathy were positively and uniquely associated with self-reported prosocial behaviour. Cognitive reappraisal, but not expressive suppression, played a role in moderating the association between empathy and prosocial behaviour. Specifically, level of cognitive reappraisal moderated the relationship between affective empathy and prosocial behaviour.

The finding that both affective and cognitive empathy are associated with prosocial behaviour supports previous studies

Table 1. Correlations between questionnaire measures.

	QCAE: CE	QCAE: AE	PTM total	ERQ: reappraisal
QCAE: AE	.417**			
PTM total	.433**	.358**		
ERQ: reappraisal	.333**	.173	.113	
ERQ: suppression	−.360**	−.529**	−.266**	−.089

Abbreviations: QCAE-AE, Questionnaire of Cognitive and Affective Empathy Affective Empathy subscale; QCAE-CE, Questionnaire of Cognitive and Affective Empathy Cognitive Empathy subscale; ERQ, Emotion Regulation Questionnaire; PTM, Prosocial Tendencies Measure.
*p<.05.
**p<.01.
doi:10.1371/journal.pone.0096555.t001

suggesting that empathy is a key motivating factor for prosocial behaviour (e.g. [3], [8] [10], [12], [15]). Interestingly, associations between affective and cognitive empathy and prosocial behaviour were not significantly different. Additional analyses showed that cognitive and affective empathy uniquely predicted prosocial behaviour, suggesting that both empathy components play a role in motivating prosocial behaviour. Consequently, whilst it is likely that these two components will often work together in everyday life as they are moderately correlated (e.g. [27], [34]), our finding raises the possibility that having high levels of just one component

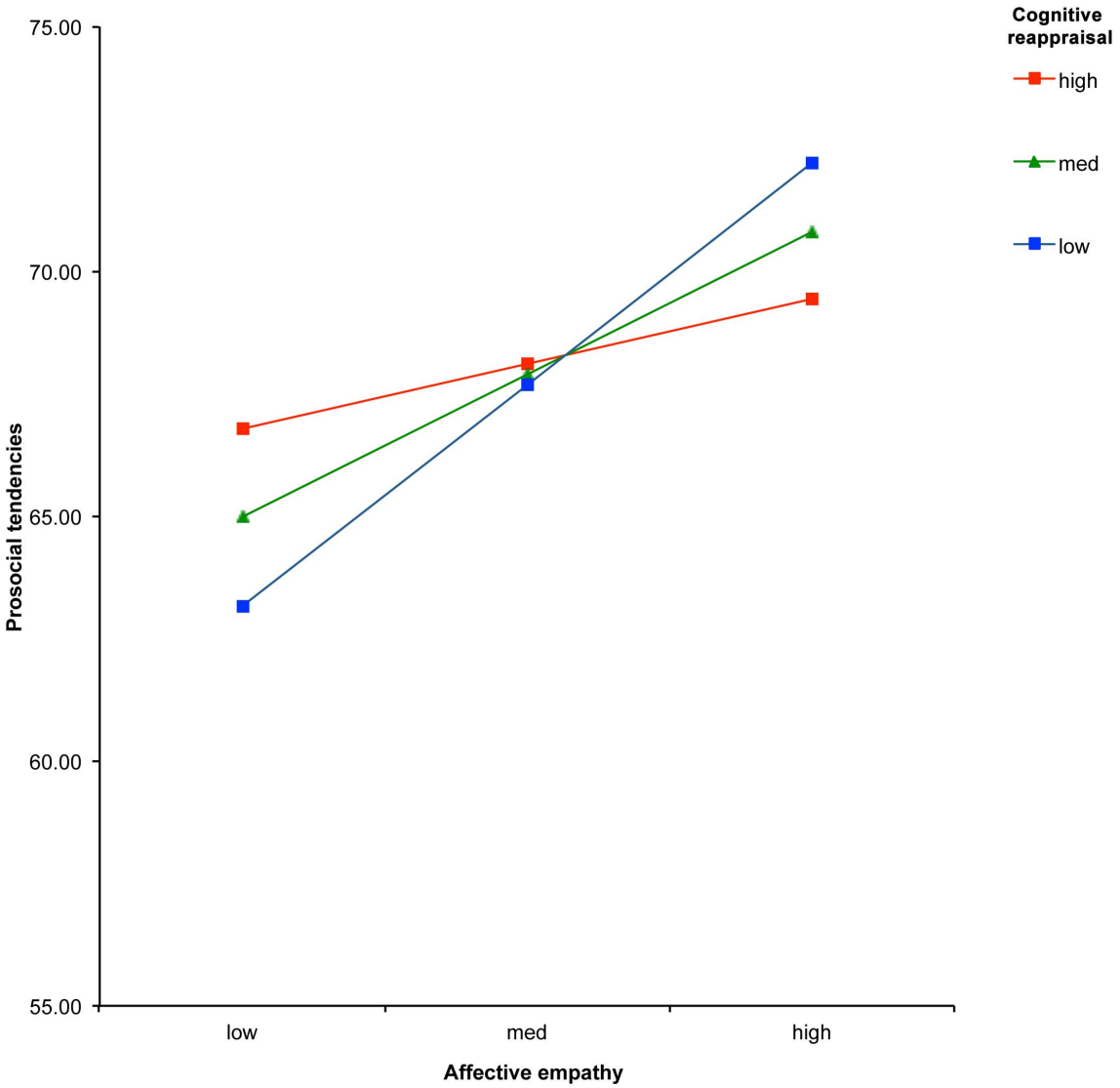


Figure 1. Moderation of the association between affective empathy and prosocial tendencies by cognitive reappraisal.
doi:10.1371/journal.pone.0096555.g001

could motivate prosocial behaviour, but this needs to be investigated further.

We also observed that expressive suppression was negatively associated with prosocial tendencies. This pattern fits with previous studies suggesting that expressive suppression is a maladaptive emotion regulation strategy [16], [24–26]. Our results extend these findings by suggesting that in, addition to being related to greater experience of negative emotions, avoidance of close relationships and reports of less life satisfaction [16], [24–26], expressive suppression is also associated with less self-reported prosocial tendencies.

The type of emotion regulation strategy was important for moderating associations between empathy and prosocial tendencies; cognitive reappraisal moderated associations between affective empathy and prosocial behaviour whilst expression suppression did not. In addition, the degree of emotion regulation interacted with the degree of empathy to predict prosocial behaviour. Affective empathy was positively associated with prosocial behaviour for participants at low and average levels of cognitive reappraisal. This positive association was not evident in participants who reported a high tendency to reappraise. Instead, these individuals had similar levels of prosocial tendencies regardless of level of affective empathy.

Consequently, although empathy is generally assumed to have a significant positive association with prosocial behaviour [3], [4] this may not be the case for all aspects of empathic processing. Our finding suggests that affective empathy is an important motivating factor for prosocial behaviour only for particular individuals, which fits with accounts considering a multitude of factors involved in motivating prosocial behaviour [5]. One explanation is that those with high tendency to reappraise are (at least according to their self-report) more able to change their strategy and viewpoint when evaluating the situation at hand. This capacity may allow one to more readily deduce the desirability of prosocial behaviour even without the experience of the affective components empathy. Whilst we observed a significant moderation of cognitive reappraisal on the association between affective empathy and prosocial behaviour, moderation effects were not evident for associations between cognitive empathy and prosocial behaviour. This lack of association could be because of the overlap in processes involved in cognitive empathy and those involved in cognitive reappraisal. Indeed self-reports of cognitive empathy and cognitive reappraisal were positively correlated in this sample. Processes such as shifting perspective or attention are common to both cognitive empathy and reappraisal. In terms of increasing prosocial behaviour in those individuals high in reappraisal, it is possible that promoting cognitive empathy might elevate the motivation of these individuals to behave prosocially.

Interestingly, we also found that those with the highest levels of self-reported prosocial behaviour were individuals low in reappraisal but high in affective empathy. Given that cognitive reappraisal is positively related to interpersonal functioning [16], [24–26] and prosocial behaviour is generally seen as a positive aspect of interpersonal functioning this result may seem somewhat surprising. In addition, the model proposed by Eisenberg & Fabes [14] suggests that those high in experiences of emotional intensity and low in emotion regulation would not manage appropriate prosocial responding and might even display antisocial/aggressive behaviour in response to emotional arousal. However, it has been suggested that high levels of prosocial and altruistic behaviour are not always beneficial and there are cases when acts that are subjectively prosocial can be, to the observer, objectively unhelpful [35]. Future research needs to determine whether the self-reported prosocial behaviours by individuals with high affective empathy and low cognitive appraisal capacities are perceived as objectively helpful/prosocial by the observer. Items on the prosocial tendencies questionnaire assess the self-reported tendency to engage in prosocial behaviours, rather than the quality of them. Future studies could include experimental and/or observational measures to examine this. The types of prosocial responses of individuals high in affective empathy and low in cognitive reappraisal could be compared to those high in cognitive reappraisal and high in affective empathy. Another promising avenue for future research is to investigate empathy components and emotion regulation strategies in tandem in clinical populations thought to be characterised by atypical empathy and emotion regulation. For example, autism spectrum disorders, psychopathy and alexithymia have all been associated with both atypical empathy and emotion regulation [36], [37]. Finally, the role of empathic concern, i.e. sympathy, in motivating prosocial behaviour has recently been studied theoretically by mathematical models [38], [39]. These models suggest that the development of empathic concern can lead to development of cooperation in economic games (termed evolutionary games by the authors). Consequently, such models suggest potential mathematical principles that could be applied in future studies to model how empathy might lead to prosocial behaviour. In parallel, our findings also suggest the potential inclusion of parameters indexing emotion regulation strategies in future models as an avenue of further research.

Conclusion

Overall, our findings suggest that both affective and cognitive empathy are motivating factors for prosocial behaviour. However, empathy and emotion regulation can also interact to predict different levels of self-reported prosocial behaviour such that there is not always a significant positive association between affective empathy and prosocial behaviour. Our results could help to account for why associations between empathy and prosocial behaviour can sometimes be modest or weak. Our results also suggest that further investigations of the type of prosocial behaviours exhibited by individuals with varying levels of empathy and emotion regulation could be relevant as we try to understand how empathy might motivate prosocial ways of interacting with others.

Acknowledgments

We would like to thank Miss Madeleine Bridge for her help with data collection.

Author Contributions

Conceived and designed the experiments: PLL EV. Performed the experiments: PLL. Analyzed the data: PLL ASC EV. Wrote the paper: PLL ASC EV.

References

1. Fehr E, Fischbacher U (2003) The nature of human altruism. Nature 425: 785–791.
2. Singer T, Lamm C (2009) The social neuroscience of empathy. Ann N Y Acad Sci 1156: 81–96.
3. Eisenberg N, Eggum ND, Di Giunta L (2010) Empathy-Related Responding: Associations with Prosocial Behavior, Aggression, and Intergroup Relations. Soc Issues Policy Rev 4: 143–180.

4. Hoffman ML (2008) Empathy and prosocial behavior. In: Lewis M, Haviland-Jones JM, Barrett LF, editors. Handbook of Emotions, Third Edition. Guilford Publications. 440–455.
5. Penner LA, Dovidio JF, Piliavin JA, Schroeder DA (2005) Prosocial behavior: Multilevel perspectives. Annu Rev Psychol 56: 365–392.
6. Davis MH (1980) A multidimensional approach to individual differences in empathy. JSAS Cat. Select. Docs Psychol 10: 85.
7. Eisenberg N, McCreath H, Ahn R (1988) Vicarious Emotional Responsiveness and Prosocial Behavior Their Interrelations in Young Children. Pers Soc Psychol Bull 14: 298–311.
8. Eisenberg N, Fabes RA, Miller PA, Fultz J, Shell R, et al. (1989) Relation of sympathy and personal distress to prosocial behavior: a multimethod study. J Pers Soc Psychol 57: 55–66.
9. Davis MH (1983) Measuring individual differences in empathy: Evidence for a multidimensional approach. J Pers Soc Psychol 44: 113–126.
10. Batson C (1998) Prosocial behavior and altruism. In: Lindzey G, Gilbert D, Fiske ST, editors. The Handbook of Social Psychology, Oxford University Press, 282–316.
11. Carlo G, Allen JB, Buhman DC (1999) Facilitating and disinhibiting prosocial behaviors: The nonlinear interaction of trait perspective taking and trait personal distress on volunteering. Basic Appl Soc Psych 21: 189–197.
12. Carlo G, Hausmann A, Christiansen S, Randall BA (2003) Sociocognitive and behavioral correlates of a measure of prosocial tendencies for adolescents. J Early Adolesc 23: 107–134.
13. Eisenberg N (2000) Emotion, regulation, and moral development. Annu Rev Psychol 51: 665–697.
14. Eisenberg N, Fabes RA (1992) Emotion, regulation, and the development of social competence. In: Clark MS, editors. Review of Personality and Social Psychology: Emotion and Social Behaviour. Newbury Park, CA: Sage. 119–150.
15. Hoffman ML (2001) Empathy and Moral Development: Implications for Caring and Justice: Cambridge University Press.
16. Gross JJ (2001) Emotion regulation in adulthood: Timing is everything. Current Directions in Psychological Science 10: 214–219.
17. Gross JJ, Thompson RA (2007) Emotion regulation: Conceptual foundations. In: Gross JJ, editors. Handbook of Emotion Regulation, The Guildford Press 3–27.
18. Gross JJ (2013) Emotion regulation: Taking stock and moving forward. Emotion 13: 359–365.
19. Gross JJ, John OP (2003) Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. J Pers Soc Psychol 85: 348–362.
20. Ray RD, Ochsner KN, Cooper JC, Robertson ER, Gabrieli JD, et al. (2005) Individual differences in trait rumination and the neural systems supporting cognitive reappraisal. Cogn Affect Behav Neurosci 5: 156–168.
21. Richards JM, Gross JJ (2000) Emotion regulation and memory: the cognitive costs of keeping one's cool. J Pers Soc Psychol 79: 410–424.
22. Gross JJ (1998) Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. J Pers Soc Psychol 74: 224–237.
23. Srivastava S, Tamir M, McGonigal KM, John OP, Gross JJ (2009) The social costs of emotional suppression: a prospective study of the transition to college. J Pers Soc Psychol 96: 883–897.
24. Gross JJ, Levenson RW (1993) Emotional suppression: physiology, self-report, and expressive behavior. J Pers Soc Psychol 64: 970–986.
25. John OP, Gross JJ (2004) Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development. J Pers 72: 1301–1334.
26. Soto JA, Perez CR, Kim YH, Lee EA, Minnick MR (2011) Is expressive suppression always associated with poorer psychological functioning? A cross-cultural comparison between European Americans and Hong Kong Chinese. Emotion 11: 1450–1455.
27. Reniers RL, Corcoran R, Drake R, Shryane NM, Vollm BA (2011) The QCAE: a Questionnaire of Cognitive and Affective Empathy. J Pers Assess 93: 84–95.
28. Baron-Cohen S, Richler J, Bisarya D, Gurunathan N, Wheelwright S (2003) The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. Philos Trans R Soc Lond B Biol Sci. 358: 361–374.
29. Hogan R (1969) Development of an empathy scale. J Consult Clin Psychol 33: 307–316.
30. Eysenck SB, Eysenck HJ (1978) Impulsiveness and venturesomeness: Their position in a dimensional system of personality description. Psychol Rep 43: 1247–1255.
31. Carlo G, Randall BA (2002) The development of a measure of prosocial behaviors for late adolescents. J Youth Adolesc 31: 31–44.
32. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 289–300.
33. Hayes AF (2013) Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-Based Approach: Guilford Press.
34. Lockwood PL, Bird G, Bridge M, Viding E (2013) Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. Front Human Neurosci 7: 760.
35. Oakley BA (2013) Concepts and implications of altruism bias and pathological altruism. Proc Natl Acad Sci 110: 10408–10415.
36. Schipper M, Petermann F (2013) Relating empathy and emotion regulation: Do deficits in empathy trigger emotion dysregulation? Soc Neurosci 8: 101–107.
37. Swart M, Kortekaas R, Aleman A (2009) Dealing with feelings: Characterization of trait alexithymia on emotion regulation strategies and cognitive-emotional processing. PLoS One, 4: e5751.
38. Szolnoki A, Xie NG, Wang C, Perc M (2011) Imitating emotions instead of strategies in spatial games elevates social welfare. Europhys Lett 96: 38002.
39. Szolnoki A, Xie NG, Ye Y, Perc M (2013) Evolution of emotions on networks leads to the evolution of cooperation in social dilemmas. Phys Rev E 87: 042805.

Current Biology 23, 901–905, May 20, 2013 ©2013 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2013.04.018

Report

Association of Callous Traits with Reduced Neural Response to Others' Pain in Children with Conduct Problems

Patricia L. Lockwood,[1,5] Catherine L. Sebastian,[1,2,5] Eamon J. McCrory,[1] Zoe H. Hyde,[1] Xiaosi Gu,[3] Stéphane A. De Brito,[4] and Essi Viding[1,*]
[1]Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London WC1H 0AP, UK
[2]Department of Psychology, Royal Holloway, University of London, Egham Hill, Egham, Surrey TW20 0EX, UK
[3]Wellcome Trust Centre for Neuroimaging, University College London, Queen Square, London WC1N 3BG, UK
[4]School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Summary

Children with conduct problems (CP) persistently violate others' rights and represent a considerable societal cost [1]. These children also display atypical empathic responses to others' distress [2], which may partly account for their violent and antisocial behavior. Callous traits index lack of empathy in these children and confer risk for adult psychopathy [3]. Investigating neural responses to others' pain is an ecologically valid method to probe empathic processing [4], but studies in children with CP have been inconclusive [5, 6]. Using functional magnetic resonance imaging (fMRI), we measured neural responses to pictures of others in pain (versus no pain) in a large sample of children with CP and matched controls. Relative to controls, children with CP showed reduced blood oxygen level-dependent responses to others' pain in bilateral anterior insula (AI), anterior cingulate cortex (ACC), and inferior frontal gyrus, regions associated with empathy for pain in previous studies [7, 8]. In the CP group, callous traits were negatively associated with responses to others' pain in AI and ACC. We conclude that children with CP have atypical neural responses to others' pain. The negative association between callous traits and AI/ACC response could reflect an early neurobiological marker indexing risk for empathic deficits seen in adult psychopathy.

Results

Conduct problems (CP) in children include aggression, theft, and cruelty to others [9]. Children with CP are considerably more likely to engage in antisocial behavior in adulthood than typically developing children and are at risk for developing adult psychopathy [3]. Antisocial behaviors displayed by children with CP may reflect atypical empathic responses to others' suffering [2]. Empathy is the capacity to understand and resonate with the affective experience of another [10] and plays a key role in inhibiting aggression and promoting prosocial behavior [11, 12]. Callous and unemotional (CU) traits index low empathy in children with CP, as well as diminished guilt, callous use of others, and shallow emotions [13, 14].

One method for investigating neural processing of empathy is to measure responses to others' pain [4]. Delineating these responses in children with CP is of particular interest because this group often inflicts pain on others [1]. fMRI studies in healthy populations have identified a network of brain regions activated by both the experience and observation of pain. This network includes sensory regions such as somatosensory cortex, affective-motivational regions (linked to processing emotional responses to pain), such as anterior insula (AI) and anterior cingulate cortex (ACC), and cognitive-regulatory regions, such as inferior frontal gyrus (IFG) [7, 8, 10, 15, 16].

Atypical function and structure in several of these regions, including AI, ACC, and prefrontal cortex, have been implicated in the pathophysiology of childhood CP and adult psychopathy [17, 18]. However, to date, only two studies have investigated neural processing of empathic pain in children with CP [5, 6], with inconclusive results. Decety et al. [6] found that, compared with controls, children with CP showed increased neural responses to others in pain in regions including the insula, anterior midcingulate, striatum, and amygdala. Aggressive CP symptoms were positively correlated with IFG, cingulate cortex, amygdala, and periaqueductal gray responses. However, CU traits were not measured, and the CP sample was small (n = 8), making replication and extension of this work important. Another study measured event-related potentials and found reduced responses to others' pain in children with CP and high levels of CU traits [5]. These findings provide preliminary evidence that children with CP show atypical responses to others' pain, which may be partially driven by CU traits.

To test the hypothesis that children with CP would show atypical neural responses to others' pain, we recorded fMRI responses to pictures of others' hands and feet either in pain or in no pain (from [19]) in a large sample of children with CP (n = 37) and controls (n = 18). Groups were matched for IQ, age, socioeconomic status, and ethnicity. Participants performed an incidental hand/foot judgment task to ensure they were attending to the stimuli. We also acquired parent and teacher ratings of CU traits using the Inventory of Callous-Unemotional Traits (ICU) [20], a standard research measure comprising callous, uncaring, and unemotional subscales. On the basis of previous research suggesting reduced empathy in children with CP [2, 13, 14], we predicted reduced neural responses in three regions of interest (ROI): AI, ACC, and IFG, all linked to empathy for pain in previous studies [7, 8, 10]. We further predicted that callous traits would be negatively associated with AI and ACC response, because response in these regions has been related to affective aspects of empathy and callous traits in particular index poor empathy.

Results from whole-brain analyses for the main effect of Pain > No Pain (and the reverse) and the group by condition interaction are displayed in Table S1 available online (see also Figure S1 and Supplemental Discussion). Main effects were found in regions previously associated with empathy for pain and largely replicated a previous study using the same stimuli [19]. ROI analyses for Pain > No Pain revealed the predicted pattern of reduced response in the CP relative to control

[5]These authors contributed equally to this work
*Correspondence: e.viding@ucl.ac.uk

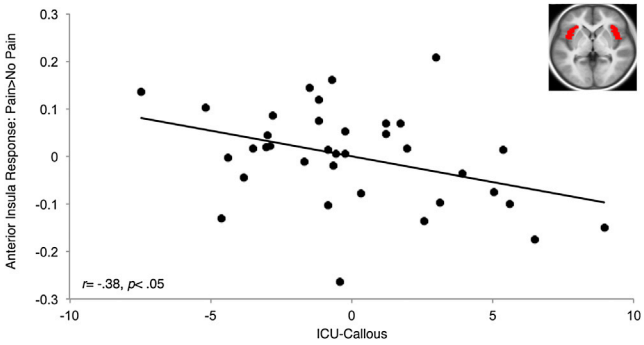


Figure 1. Partial Regression Plot for the CP Group Showing a Negative Relationship between Bilateral AI Response to Pain > No Pain and Unique Variance Associated with ICU-Callous Traits

Partial regression plot for the CP group (n = 36) shows a negative relationship between bilateral AI response to Pain > No Pain and unique variance associated with ICU-callous traits after controlling for CASI-CD, ICU-unemotional, and ICU-uncaring scores. Inset shows horizontal section (z = 0) of bilateral AI ROI overlaid on an average T1 structural image from all participants. Bilateral AI response was calculated by averaging left and right AI response. *P* and r reflect partial correlation coefficients.

group in bilateral AI (t[53] = 2.08, p = 0.02), ACC (t[53] = 1.66, p = 0.05), and IFG (t[53] = 2.45, p < 0.01) (see Experimental Procedures and Supplementary Experimental Procedures for full details of analyses, including ROI definition and statistical thresholds). Levene's test indicated that variance did not differ between groups for any ROI (p values > 0.20).

We then examined our second hypothesis, that callous traits would be associated with reduced ROI responses to Pain > No Pain within the CP group. On the basis of previous findings showing that CP symptoms and CU traits exert suppressor effects on one another (see [21, 22]), we conducted multiple regressions to investigate unique contributions of ICU subscales (callous, uncaring, unemotional) and CP symptoms to neural responses in our ROIs (see Table S2 for bivariate correlations). One participant was excluded from these analyses due to missing data.

In AI, a significant negative relationship was found between unique variance associated with callous traits and neural response (β = −0.625, p = 0.029) (Figure 1). Neither CP symptoms nor uncaring or unemotional subscales were associated with AI response (all p values > 0.10). In ACC, a significant negative relationship was found between unique variance associated with callous traits and neural response (β = −0.729, p = 0.010), whereas a positive relationship was found between unique variance associated with CP symptom scores and neural response (β = 0.485, p = 0.019) (Figure 2). No relationships were found in relation to the uncaring or unemotional subscale scores (p values > 0.24). In IFG, no associations with unique variance were found. To investigate potential effects of commonly comorbid attention-deficit hyperactivity, generalized anxiety, and depression symptoms, we included these variables in follow-up regression analyses. All significant results remained at p < 0.05, and none of these variables predicted AI or ACC response (all p values > 0.25). In addition, when age was included in follow-up regression analyses, all results remained significant at p < 0.05.

Behavioral data from the incidental hand/foot judgment task showed a main effect of condition for both reaction times (F[1,53] = 71.85, p < 0.001) and errors (F[1,53] = 6.40, p = 0.014), with significantly slower RTs and greater error rates in the pain condition (mean RT = 910 ms, SD = 140 ms; mean % error = 6.82, SD = 5.05) compared with no pain (mean RT = 863 ms, SD = 130 ms; mean % error = 5.63, SD = 4.55). We therefore reran all regression models, controlling for RT and error-rate difference scores (pain − no pain) to exclude the possibility that differing cognitive conflict demands could account for our findings. All results remained significant at p < 0.05. RT and error data showed no main effects of group or interaction between group and condition (see section "Behavioral Data" in Supplemental Results).

Discussion

We demonstrate reduced neural responses to others' pain in children with conduct problems compared with matched controls in three regions (bilateral AI, ACC, and IFG) associated with affective-motivational and cognitive-regulatory empathic processing. This is the first fMRI study to investigate empathy for pain processing in a large sample of children with CP, using a well-controlled task matched for visual and social content. We also show a negative association between callous traits and responses in AI and ACC, regions related to unpleasant emotions generated in response to others' pain [7, 8].

Meta-analyses indicate that AI and ACC are consistently activated during empathy for pain and have a close functional relationship [7, 8, 16]. AI plays an important role in sensory integration [23] and interoceptive awareness [24] and may be involved in awareness of unpleasant feelings during empathy for pain [16]. Interestingly, abnormal AI function and structure have frequently been reported in both children with CP and in adults with psychopathy [6, 21, 25, 26]. However, our observation of reduced AI response is at odds with one study [6], which found increased AI response in children with CP. This could be because in that study [6], pain caused by accident was contrasted with pain caused by others, whereas our pain and no-pain conditions were matched for agency. Increased AI reactivity [6] may reflect differences in agency processing rather than pain processing per se. Differences in the samples between the two studies (e.g., levels of callous traits) may also have contributed to the divergent findings. Our data provide additional support for the view that atypical AI function represents a neural marker of disrupted empathic processing in CP and that AI hypoactivity relates to differences in processing others' pain.

It has been suggested that ACC mediates responses to unpleasant negative emotion generated in AI [16]. However, the role of ACC in empathy may be more domain general than that of AI, given its involvement in general information processing [19, 27]. Like AI, atypical ACC function in CP has been reported previously, again with mixed findings [6, 28]. One study reported reduced ACC response to negative pictures in CP [28], whereas another found greater ACC response in children with CP to videos of others in pain versus no pain [6]. Our finding provides converging evidence that ACC function is atypical in CP and in particular that there is hypoactivity of response during empathy for pain.

The pattern of reduced IFG response is of interest, given the known involvement of this region in emotion regulation and pain suppression [15, 29, 30] as well as in empathy tasks [7, 8]. It is possible that fewer regulatory resources were required, given that responses in other regions processing empathy for pain were hypoactive. It could also be that the result reflects known deficiencies in emotion regulation in children with CP [31].

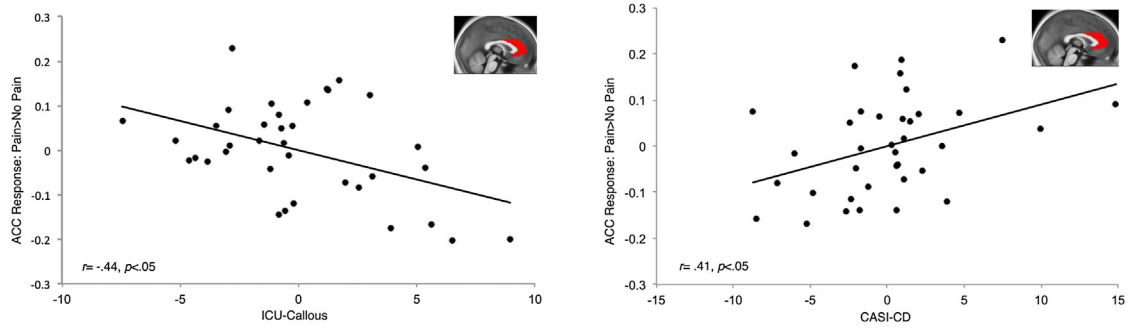


Figure 2. Partial Regression Plots Showing Associations with ACC Response to Pain > No Pain in the CP Group

Partial regression plots showing associations with ACC response to Pain > No Pain in the CP group (n = 36). Left: negative relationship between ACC response and unique variance associated with ICU-callous traits, after controlling for CASI-CD, ICU-unemotional, and ICU-uncaring scores. Right: positive relationship between ACC response and unique variance associated with CASI-CD scores, after controlling for ICU-callous, unemotional, and uncaring subscale scores. Insets: sagittal section (x = 0) of ACC ROI overlaid on an average T1 structural image from all participants. Bilateral ACC response was calculated by averaging left and right ACC response. *P* and r reflect partial correlation coefficients.

To address our second aim, we explored dimensional contributions of CU traits and CP symptoms to ROI responses. As predicted, unique variance associated with callous traits was negatively related to AI and ACC response. Because the callous subscale of the ICU contains items reflecting poor empathy in everyday life, our findings provide evidence of convergent validity between questionnaire and neural measures of empathy in CP. Moreover, callousness is an important feature of adult psychopathy [32], and childhood CU traits predict adult psychopathy [3]. Blunted neural responses to pain in children with higher levels of callous traits may characterize a developmental vulnerability to serious antisocial behavior; for a minority, such a pattern may interact with other vulnerability factors to increase risk of adult psychopathy.

Interestingly, CP symptoms were positively correlated with ACC response. These results complement recent findings [21] showing opposing unique contributions of CU traits and CP symptoms to neural response in the amygdala. Heterogeneity in CP may help to explain inconsistencies across previous studies reporting both increased and decreased ACC responses in CP [6, 28]. Importantly, differences in cognitive conflict (as indexed by RT and error differences between pain and no-pain conditions) did not account for the ACC findings. More generally, these data highlight that children with CP are a heterogeneous group with varying neurocognitive vulnerabilities, with callous traits of particular importance in predicting empathic dysfunction.

Limitations of the current study include the use of a research diagnosis of CP and a focus on males. Replication in a clinically diagnosed sample will be important, as will investigation of potential gender differences. Additionally, our task did not allow us to explore the function of component processes within the empathy for pain response in CP. Future studies should address whether there is a specific aspect of this response that is atypical in CP, e.g., basic arousal, interoceptive processing, or higher-level emotional responses to others' suffering. Finally, replication and extension of the current study is required. In particular, longitudinal studies documenting the development and persistence of reduced neural responses to others' pain in children with CP would be informative.

Despite these limitations, our data extend understanding of the neural basis of CP and empathy in several important ways. To our knowledge, this is the first study to investigate empathic pain processing in a large sample of children with CP compared with controls on a task matched for visual and social content. First, we show reduced neural responses to others' pain in children with CP. Second, we show that callous traits in particular may underlie atypical neural responses to others' pain in CP, which may represent an early neurobiological marker for later psychopathy. Third, the finding that callous traits and CP symptoms show opposing relationships with ACC response suggests a potential explanation for mixed reports of hyperactivation [6] and hypoactivation [28] of ACC to negative affective stimuli in CP. Clinically, our data may have consequences for empathy training implementation (e.g., in relation to victim empathy [33]) in children with high levels of callous traits. Systematic evaluation of training outcomes should take callous traits into account. It remains an empirical question whether empathic responding can be normalized in children with CP (and varying levels of callous traits) or whether behavioral equivalence is better achieved through compensatory strategies that leverage spared cognitive processes [13, 14].

Experimental Procedures

Participants

Right-handed boys aged 10–16 (mean [SD]: controls = 13.68 [1.68]; CPs = 14.05 [1.69]) were recruited from the community via advertisements and local schools. Screening questionnaires were completed by 143 parents and teachers. CP was assessed using the Child and Adolescent Symptom Inventory (CASI-4R) [34] Conduct Disorder scale (CASI-CD). CASI-CD cutoff scores for inclusion in the CP group were: parent report = 4+ (ages 10–12) and 3+ (ages 12–16) or teacher report = 3+ (ages 10–12), 4+ (ages 12–14), and 6+ (ages 15–16). These scores are associated with a clinical diagnosis of CD [35]. CU traits were assessed using the Inventory of Callous-Unemotional Traits (ICU) [20]. Total scores for the three ICU subscales (callous, uncaring, and unemotional) were calculated [20]. Both CASI-CD and ICU were scored by taking the highest ratings from either the parent or teacher questionnaire for each item [36]. For two children with CP, only parent ratings were available. The Strengths and Difficulties Questionnaire (SDQ [37]) was used to screen for psychopathology (hyperactivity, CP, emotional symptoms, peer problems) in the controls. All control participants scored below the CP group median on the ICU and in the normal range on the CASI-CD and SDQ. For both groups, exclusion criteria included previous diagnosis of neurological or psychotic disorder, including autism spectrum disorders, and current prescription for psychiatric medication (all children were unmedicated). Written informed consent from parents and written assent from participants was obtained.

Table 1. Participant Characteristics

Characteristics and Questionnaires	Group		
	Controls (n = 18)	CP (n = 37)	p Value[a]
Demographic Variables			
Age[b]	13.68 (1.68)	14.05 (1.69)	0.456
Socioeconomic Status[b]	3.07 (1.01)	3.23 (1.26)	0.635
F-IQ[c,d]	102.83 (11.69)	101.17 (13.46)	0.656
V-IQ[c,d]	53.06 (8.73)	49.92 (10.96)	0.295
P-IQ[c,d]	49.67 (8.61)	50.33 (9.57)	0.804
Ethnicity[b,e]	13:2:2:1	26:3:6:2	1.00
ICU[f]	24.17 (4.85)	42.97 (10.67)	<0.001
Child and Adolescent Symptom Inventory			
Conduct Disorder[f]	0.56 (0.70)	10.14 (6.18)	<0.001
ADHD[d,g]	9.47 (7.47)	25.04 (13.75)	<0.001
Generalized Anxiety Disorder[d,g]	2.71 (3.07)	7.46 (5.37)	0.001
Major Depressive Episode[d,g]	2.61 (1.09)	6.38 (5.40)	0.005
AUDIT[c,d]	1.22 (1.99)	2.11 (3.88)	0.366
DUDIT[c,d,h]	0.17 (0.51)	2.50 (6.62)	0.143

Abbreviations: CP, conduct problems; F-IQ, full IQ score from the two-subtest Wechsler Abbreviated Scale of Intelligence; V-IQ, verbal IQ score; P-IQ, matrix reasoning IQ score; ICU, Inventory of Callous-Unemotional Traits; ADHD, attention-deficit/hyperactivity disorder; AUDIT/DUDIT, Alcohol/Drug Use Disorders Identification Test.
[a]All p values were obtained using t tests except for ethnicity (Fisher's exact test used).
[b]Measures taken at screening phase—parent report.
[c]Child at scanning session.
[d]Missing data from 1 CP.
[e]Ethnicity: White:Black:Mixed:Asian.
[f]Meaures taken at screening phase—parent and teacher report.
[g]Measures taken at scanning session—parent report.
[h]The Drug Use Disorders Identification Test requires participants to rate the frequency of any substance use on a five-point scale from "never" to "almost daily." The list of drugs includes cannabis, amphetamines, cocaine, opiates, hallucinogens, solvents, and GHB, as well as medicines used in an abusive way.

Fifty-eight children were scanned (39 CPs, 19 controls), with usable data from 37 CPs and 18 controls. Exclusions were due to scanner refusal (1 CP) and teacher questionnaire data obtained after scanning indicating that the child no longer met group criteria (1 CP, 1 control). Groups were matched on IQ, age, ethnicity, and socioeconomic status (Table 1).

Experimental Task

Stimuli were 192 digital photographs showing another person's hand or foot in painful or nonpainful situations [19]. "Pain" and "No Pain" stimuli (96 pictures per condition) were matched on physical properties and were validated as eliciting empathy-related activations in a previous study [19]. Stimuli were presented in pain and no-pain blocks lasting 20 s and consisting of eight images, each displayed for 2,000 ms with a 500 ms interstimulus interval. Blocks were pseudorandomized, with the same block type never presented more than twice in a row. A fixation cross was presented for 15 s every six blocks.

To ensure attention, participants performed a hand/foot key press judgment on every trial. Participants practiced outside the scanner with painful and nonpainful images not seen in the main experiment, until $\geq$80% accuracy was reached.

Psychometric and Questionnaire Measures

Participants completed the Wechsler Abbreviated Scale of Intelligence two-subtest version [38] and the Alcohol and Drug Use Disorders Identification Tests [39, 40]. A parent or guardian completed the CASI-4R scales for symptoms commonly comorbid with CP, including ADHD, generalized anxiety disorder, and major depressive episode (Table 1).

fMRI Data Acquisition and Analysis

A Siemens Avanto 1.5 T MRI scanner was used to acquire 189 multislice T2*-weighted echo planar volumes with blood oxygenation level-dependent contrast (one run of 9 min). The sequence was based on Weiskopf et al. [41] (see Supplemental Information for acquisition parameters, preprocessing pipeline, and procedures for removing data corrupted by participant motion). A 5.5 min T1-weighted MPRAGE scan was acquired for coregistration, normalization, and overlay, and fieldmaps were acquired for unwarping. Data were analyzed using Statistical Parametric Mapping (SPM8; http://www.fil.ion.ucl.ac.uk/spm).

After standard preprocessing, a block analysis compared neural activity associated with pain and no-pain conditions. Regressors included Pain and No Pain (blocks of 20 s duration) and fixation (15 s), modeled as boxcar functions convolved with a canonical hemodynamic response function. The six realignment parameters were also modeled as effects of no interest.

At the first level, Pain > No Pain and No Pain > Pain contrasts were created. Contrast images were entered into second-level analyses, where group (CP versus control) served as a between-subjects variable in independent sample t tests. Main effects are reported at $p < 0.05$, family-wise error (FWE) corrected across the whole brain, whereas regions from a whole-brain analysis showing a condition by group interaction are presented at $p < 0.005$, $k \geq 10$, uncorrected (no interaction results survived FWE correction across the whole brain) (Table S1). We investigated the condition by group interaction in three a priori regions of interest (bilateral AI, ACC, and IFG). ROIs were anatomically defined using masks from the automated anatomical labeling atlas [42]. The MarsBaR toolbox (http://marsbar.sourceforge.net) was used to calculate mean contrast estimates across bilateral ROIs. Group differences were assessed at a standard statistical threshold of $p < 0.05$ [43, 44].

Supplemental Information

Supplemental Information includes one figure, two tables, Supplemental Results, Supplemental Discussion, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.04.018.

Acknowledgments

This work was supported by award 53229 from the British Academy, award RES-062-23-2202 from the Economic and Social Research Council (both to E.V. [principal investigator] and E.J.M.), and funding from the Birkbeck–University College London Centre for Neuroimaging. P.L.L. was supported by a Doctoral Training Account studentship from the Medical Research Council. The authors thank Oliver Josephs for technical advice, Philip Kelly and Elizabeth O'Nions for help with data collection, and all the children, parents, teachers, and schools who took part in this study. All aspects of the study were approved by the University College London research ethics committee (project identifier: 0622/001).

Received: January 2, 2013
Revised: March 14, 2013
Accepted: April 8, 2013
Published: May 2, 2013

References

1. Romeo, R., Knapp, M., and Scott, S. (2006). Economic cost of severe antisocial behaviour in children—and who pays it. Br. J. Psychiatry *188*, 547–553.
2. Blair, R.J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. Conscious. Cogn. *14*, 698–718.
3. Lynam, D.R., Caspi, A., Moffitt, T.E., Loeber, R., and Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. J. Abnorm. Psychol. *116*, 155–165.
4. Jackson, P.L., Meltzoff, A.N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. Neuroimage *24*, 771–779.
5. Cheng, Y., Hung, A.Y., and Decety, J. (2012). Dissociation between affective sharing and emotion understanding in juvenile psychopaths. Dev. Psychopathol. *24*, 623–636.
6. Decety, J., Michalska, K.J., Akitsuki, Y., and Lahey, B.B. (2009). Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. Biol. Psychol. *80*, 203–211.

7. Fan, Y., Duncan, N.W., de Greck, M., and Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. Neurosci. Biobehav. Rev. *35*, 903–911.
8. Lamm, C., Decety, J., and Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. Neuroimage *54*, 2492–2502.
9. American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR (Washington, DC, USA: American Psychiatric Association).
10. Singer, T., and Lamm, C. (2009). The social neuroscience of empathy. Ann. N Y Acad. Sci. *1156*, 81–96.
11. Eisenberg, N., Cumberland, A., Guthrie, I.K., Murphy, B.C., and Shepard, S.A. (2005). Age Changes in Prosocial Responding and Moral Reasoning in Adolescence and Early Adulthood. J. Res. Adolesc. *15*, 235–260.
12. Decety, J., and Jackson, P.L. (2004). The functional architecture of human empathy. Behav. Cogn. Neurosci. Rev. *3*, 71–100.
13. Jones, A.P., Happé, F.G.E., Gilbert, F., Burnett, S., and Viding, E. (2010). Feeling, caring, knowing: different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. J. Child Psychol. Psychiatry *51*, 1188–1197.
14. Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., Romanos, M., Schecklmann, M., Schneider, W., Warnke, A., and Freitag, C.M. (2012). Empathy in children with autism and conduct disorder: group-specific profiles and developmental aspects. J. Child Psychol. Psychiatry *53*, 651–659.
15. Ochsner, K.N., and Gross, J.J. (2005). The cognitive control of emotion. Trends Cogn. Sci. *9*, 242–249.
16. Medford, N., and Critchley, H.D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. Brain Struct. Funct. *214*, 535–549.
17. Anderson, N.E., and Kiehl, K.A. (2012). The psychopath magnetized: insights from brain imaging. Trends Cogn. Sci. *16*, 52–60.
18. Sterzer, P., and Stadler, C. (2009). Neuroimaging of aggressive and violent behaviour in children and adolescents. Front Behav Neurosci. *3*, 35.
19. Gu, X., Liu, X., Guise, K.G., Naidich, T.P., Hof, P.R., and Fan, J. (2010). Functional dissociation of the frontoinsular and anterior cingulate cortices in empathy for pain. J. Neurosci. *30*, 3739–3744.
20. Essau, C.A., Sasagawa, S., and Frick, P.J. (2006). Callous-unemotional traits in a community sample of adolescents. Assessment *13*, 454–469.
21. Sebastian, C.L., McCrory, E.J.P., Cecil, C.A.M., Lockwood, P.L., De Brito, S.A., Fontaine, N.M.G., and Viding, E. (2012). Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. Arch. Gen. Psychiatry *69*, 814–822.
22. Hicks, B.M., and Patrick, C.J. (2006). Psychopathy and negative emotionality: analyses of suppressor effects reveal distinct relations with emotional distress, fearfulness, and anger-hostility. J. Abnorm. Psychol. *115*, 276–287.
23. Critchley, H.D., Wiens, S., Rotshtein, P., Ohman, A., and Dolan, R.J. (2004). Neural systems supporting interoceptive awareness. Nat. Neurosci. *7*, 189–195.
24. Craig, A.D.B. (2009). How do you feel—now? The anterior insula and human awareness. Nat. Rev. Neurosci. *10*, 59–70.
25. Sterzer, P., Stadler, C., Poustka, F., and Kleinschmidt, A. (2007). A structural neural deficit in adolescents with conduct disorder and its association with lack of empathy. Neuroimage *37*, 335–342.
26. Ly, M., Motzkin, J.C., Philippi, C.L., Kirk, G.R., Newman, J.P., Kiehl, K.A., and Koenigs, M. (2012). Cortical thinning in psychopathy. Am. J. Psychiatry *169*, 743–749.
27. Gu, X., Gao, Z., Wang, X., Liu, X., Knight, R.T., Hof, P.R., and Fan, J. (2012). Anterior insular cortex is necessary for empathetic pain perception. Brain *135*, 2726–2735.
28. Sterzer, P., Stadler, C., Krebs, A., Kleinschmidt, A., and Poustka, F. (2005). Abnormal neural responses to emotional visual stimuli in adolescents with conduct disorder. Biol. Psychiatry *57*, 7–15.
29. Salomons, T.V., Johnstone, T., Backonja, M.-M., Shackman, A.J., and Davidson, R.J. (2007). Individual differences in the effects of perceived controllability on pain perception: critical role of the prefrontal cortex. J. Cogn. Neurosci. *19*, 993–1003.
30. Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A., and Ochsner, K.N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. Neuron *59*, 1037–1050.
31. Frick, P.J., and Morris, A.S. (2004). Temperament and developmental pathways to conduct problems. J. Clin. Child Adolesc. Psychol. *33*, 54–68.
32. Blair, J., Mitchell, D., and Blair, K. (2005). The Psychopath: Emotion and the Brain (Malden, MA, USA: Blackwell Publishing).
33. Hodgins, S., Viding, E., and Plodowski, A. (2009). The Neurobiological Basis of Violence: Science and Rehabilitation (Oxford: Oxford University Press).
34. Gadow, K.D., and Sprafkin, J. (2009). The Symptom Inventories: An Annotated Bibliography (Stony Brook, NY, USA: Checkmate Plus).
35. Gadow, K.D., and Sprafkin, J. (1998). Adolescent Symptom Inventory-4, Norms Manual (Stony Brook, NY, USA: Checkmate Plus).
36. Piacentini, J.C., Cohen, P., and Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: are complex algorithms better than simple ones? J. Abnorm. Child Psychol. *20*, 51–63.
37. Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. J. Child Psychol. Psychiatry *38*, 581–586.
38. Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence: WASI (San Antonio, TX, USA: Harcourt Assessment).
39. Santis, R., Garmendia, M.L., Acuña, G., Alvarado, M.E., and Arteaga, O. (2009). The Alcohol Use Disorders Identification Test (AUDIT) as a screening instrument for adolescents. Drug Alcohol Depend. *103*, 155–158.
40. Berman, A.H., Bergman, H., Palmstierna, T., and Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. Eur. Addict. Res. *11*, 22–31.
41. Weiskopf, N., Hutton, C., Josephs, O., and Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. Neuroimage *33*, 493–504.
42. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage *15*, 273–289.
43. Masten, C.L., Colich, N.L., Rudie, J.D., Bookheimer, S.Y., Eisenberger, N.I., and Dapretto, M. (2011). An fMRI investigation of responses to peer rejection in adolescents with autism spectrum disorders. Dev Cogn Neurosci. *1*, 260–270.
44. Eisenberger, N.I., Berkman, E.T., Inagaki, T.K., Rameson, L.T., Mashal, N.M., and Irwin, M.R. (2010). Inflammation-induced anhedonia: endotoxin reduces ventral striatum responses to reward. Biol. Psychiatry *68*, 748–754.

Behavioral/Cognitive

Encoding of Vicarious Reward Prediction in Anterior Cingulate Cortex and Relationship with Trait Empathy

Patricia. L. Lockwood,[1] Matthew A.J. Apps,[2,3] Jonathan P. Roiser,[4]* and Essi Viding[1]*

[1]Division of Psychology and Language Sciences, University College London, London WC1H 0AP, United Kingdom, [2]Nuffield Department of Clinical Neuroscience, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom, [3]Department of Experimental Psychology, University of Oxford, Oxford OX1 3DU, United Kingdom, and [4]Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom

Empathy—the capacity to understand and resonate with the experiences of others— can depend on the ability to predict when others are likely to receive rewards. However, although a plethora of research has examined the neural basis of predictions about the likelihood of receiving rewards ourselves, very little is known about the mechanisms that underpin variability in vicarious reward prediction. Human neuroimaging and nonhuman primate studies suggest that a subregion of the anterior cingulate cortex in the gyrus (ACCg) is engaged when others receive rewards. Does the ACCg show specialization for processing predictions about others' rewards and not one's own and does this specialization vary with empathic abilities? We examined hemodynamic responses in the human brain time-locked to cues that were predictive of a high or low probability of a reward either for the subject themselves or another person. We found that the ACCg robustly signaled the likelihood of a reward being delivered to another. In addition, ACCg response significantly covaried with trait emotion contagion, a necessary foundation for empathizing with other individuals. In individuals high in emotion contagion, the ACCg was specialized for processing others' rewards exclusively, but for those low in emotion contagion, this region also responded to information about the subject's own rewards. Our results are the first to show that the ACCg signals probabilistic predictions about rewards for other people and that the substantial individual variability in the degree to which the ACCg is specialized for processing others' rewards is related to trait empathy.

Key words: anterior cingulate; emotion contagion; empathy; fMRI; reward prediction; social reward

Significance Statement

Successfully cooperating, competing, or empathizing with others can depend on our ability to predict when others are going to get something rewarding. Although many studies have examined how the brain processes rewards we will get ourselves, very little is known about vicarious reward processing. Here, we show that a subregion of the anterior cingulate cortex in the gyrus (ACCg) shows a degree of specialization for processing others' versus one's own rewards. However, the degree to which the ACCg is specialized varies with people's ability to empathize with others. This new insight into how vicarious rewards are processed in the brain and vary with empathy may be key for understanding disorders of social behavior, including psychopathy and autism.

Introduction

The successful prediction of future rewards is fundamental for adaptive behavior. The neural mechanisms that underpin reward prediction for oneself are becoming increasingly well understood (Schultz, 2013). However, during social interactions, stimuli are often predictors of rewards for others, not exclusively ourselves. Effectively cooperating, competing, or empathizing with another requires the ability to compute the value of stimuli that predict rewards for others (Ruff and Fehr, 2014). However, very little is known about how vicarious reward predictions are processed in the brain. Moreover, there is a dearth of knowledge regarding

Received May 1, 2015; revised Aug. 18, 2015; accepted Aug. 23, 2015.

Author contributions: P.L.L., M.A.J.A., J.P.R., and E.V. designed research; P.L.L. performed research; P.L.L. analyzed data; P.L.L., M.A.J.A., J.P.R., and E.V. wrote the paper.

This work was supported by the Medical Research Council (doctoral training account studentship from the Medical Research Council to P.L.L.) and the Birkbeck–University College London Centre for Neuroimaging. M.A.J.A. is supported by an Anniversary Future Leaders fellowship from the Biotechnology and Biological Sciences Research Council (Grant BB/M013596/1). E.V. is supported by the Medical Research Council (Grant MR/K014080/1) and is a Royal Society Wolfson Research Merit Award holder. We thank Lewis Pollock and Jonathan Blott for acting as confederates and Philip Kelly, Elizabeth O'Nions, and Ana Seara-Cardoso for advice and help with data collection.

*J.P.R. and E.V. contributed equally to this work.

This article is freely available online through the *J Neurosci* Author Open Choice option.

Correspondence should be addressed to Patricia L. Lockwood, Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London WC1H 0AP, United Kingdom. E-mail: p.lockwood@ucl.ac.uk.

DOI:10.1523/JNEUROSCI.1703-15.2015

how individual differences in social functioning are related to neural response to others' reward.

The dorsal anterior cingulate cortex (dACC) signals predictive information about reward value, including the probability and magnitude of future rewards (Shidara and Richmond, 2002; Rogers et al., 2004; Sallet et al., 2007). This region is also engaged when processing social information (Behrens et al., 2008; Lamm et al., 2011; Gabay et al., 2014). Recently, a model of the dACC was proposed that unifies these different facets of its function (Apps et al., 2013b). This model posits that a subregion of the ACC in the gyrus (ACCg), lying in the anterior portions of the midcingulate cortex (areas 24a′/24b′; Vogt et al., 1995), is sensitive to processing information about rewards for other people, including probabilistic predictions about rewards that others are likely to receive (Apps et al., 2013b). Several lines of evidence support this model. First, there are neurons in the ACCg that respond when a monkey views cues indicating that another monkey will receive a reward (Chang et al., 2013) neurons and in the dACC that respond when monkeys predict the decisions of a conspecific in an economic game (Haroush and Williams, 2015). Second, lesions to the ACCg reduce the value assigned to social stimuli, leaving the processing of nonsocial stimuli intact (Rudebeck et al., 2006). Third, hemodynamic responses in this region vary with the net-value of rewards received by others, the volatility of social information, predictions about the value of others' actions, and predictions of social approval from others (Behrens et al., 2008; Jones et al., 2011; Boorman et al., 2013; Apps and Ramnani, 2014). Together, these studies point to a central role for the ACCg in processing information about others' rewards. However, a key untested hypothesis from this model is that the ACCg is engaged when predictions are made about the probability of another person receiving a reward. Therefore, the first aim of our study was to test this hypothesis.

A second hypothesis derived from this model is that individual differences in social functioning, specifically empathy, vary with the extent to which ACCg is specialized for processing others' rewards. Empathy can be broadly defined as the capacity to understand and resonate with the experiences of others (Singer and Lamm, 2009). Empirical and theoretical accounts have suggested that the ACC is involved in empathizing (Lamm et al., 2011; Engen and Singer, 2013), but prior work has largely focused on response of this region to processing others' pain and other negative outcomes (for review, see Lamm et al., 2011) rather than positive, rewarding outcomes. The propensity to feel empathy varies substantially between individuals (Blair, 2005; Lockwood et al., 2013a; Bird and Viding, 2014), but the mechanisms that underpin individual differences in vicariously processing another's rewards are poorly understood. Therefore, the second aim of our study was to test the hypothesis that the extent to which the ACCg is specialized for processing others' rewards is positively associated with trait empathy.

Materials and Methods

Participants

Thirty-two right-handed healthy males (age 19–32 years, M = 22.7, SD = 3.0) were recruited through university participant databases. Exclusion criteria included previous or current neurological or psychiatric disorder, non-normal or noncorrected to normal vision, non-native English language, and previous or current study of psychology. This latter criterion was used due to concerns that prior experience of studying psychology could compromise participants' belief in the deception used in the protocol. Two participants were excluded from the analysis (one due to excessive motion (>10% of scans) and one due to neurological abnormalities), leaving a final sample of 30. All participants gave written informed consent and the study was approved by the local departmental research ethics committee.

Experimental task

Design. We examined the processing of cues that signaled the probability with which a first-person and a third-person would receive a reward. A 2×2 factorial design, agency (self vs other) and probability (high 80% vs low 20%), was used to examine activation time-locked to the cues (see Fig. 1).

On each trial during the experiment, participants saw cues that indicated the probability with which they (first-person or "self") or the other participant (third-person or "other") were likely to win points. These cues were represented as pie charts to depict the level of probability explicitly and minimize any requirements for reward learning across the task. The cues for self and other differed in color, but were luminance matched. Self cues had the word "you" written above them, whereas other cues had the name of the other participant (a confederate) written above them. This ensured that participants were explicitly aware of whether the cues predicted outcomes for themselves or for the other participant.

After the cue, an outcome was presented. To ensure attention to the cues, participants indicated (at the time of the outcome) whether the outcome was expected or not with a button press. We specifically investigated passively delivered rather than instrumentally obtained rewards so that any activation differences between self and other trials could not be related to differences in motor preparation (e.g., an action on a self trial but no action on another trial).

Before scanning, participants completed a practice version of the task during which they received feedback as to whether their judgements (expected or unexpected) were correct. During scanning, however, participants were instructed that they would not receive feedback on their judgements, but that they should respond as quickly and accurately as possible to the judgment.

There were 100 trials in total, 50 self trials and 50 other trials presented in a pseudorandom order, with no more than three trials in a row of self or other cues. The 50 self trials consisted of 25 trials of high-probability first-person cues and 25 trials of low-probability first-person cues. Similarly, the 50 other trials consisted of 25 high-probability third-person cues and 25 low-probability first-person cues. For both the self and other conditions, 20 outcomes were an expected win, 20 outcomes were an expected no win, five outcomes were an unexpected win, and five outcomes were an unexpected no win (equivalent to 80%/20% probability).

Trial structure. Each trial began with a cue that signaled the probability of reward (80%/20%) and agent (self/other) for 800 ms (see Fig. 1*A*). After a jittered delay (2500–6000 ms), participants observed an outcome (win 100 points/win 0 points; 800 ms), followed by a variable fixation (2000–4000 ms). Participants were then presented with the options "yes" or "no" and were required to press one of the two buttons to indicate whether the outcome was expected or not. The side of the screen on which these options were presented was counterbalanced so that participants could not form a representation of a specific motor command at any point during a trial. Participants had 1500 ms to indicate their option or the word "missed" appeared in red on the screen. This was followed by a fixation cross (1000–2000 ms).

Procedure. Participants were paired with one of two age-matched confederates (who were also male), whom they believed were naive participants and had never met before the experiment. The participant and the confederate were instructed together that they could earn extra payment, based on the outcomes they received during the experimental task (see below); but in fact all participants were paid the same amount (total £30, representing an additional £7 to the standard participant payment for the required time commitment). They also believed that the confederate participant could earn an extra payment in the same manner during the task. A set of standardized questions completed after the scan confirmed that no participant had become suspicious about the deception during the experiment.

Participants attended two sessions. The first session was attended only by the experimental participant without a confederate and involved practicing the experimental task and completing questionnaires. In the first

session, attended only by the experimental participant, the "other" participant was described as "player 2" and the experimental participant was instructed that, in the scanning session, this name would be replaced by the name of the other participant. Participants were instructed that, during the practice session, the points would not be converted into any money either for themselves or the other person, but that when they attended the scanning session, these points would be converted into additional payment for themselves and the other participant. The second session (<7 d later) was attended by both the experimental participant and the confederate. During this session, the experimental participant performed the task while inside the MRI scanner. The experimental participant was under the impression that the confederate performed the same task simultaneously. The confederate was seated in the adjacent MRI control room to maintain this impression. The participants were instructed that, regardless of whether the cues and outcomes were for themselves or for the other person, they should perform the same judgment task to indicate whether the outcome was expected or not. Moreover, participants were not instructed to the specific payoff matrix, which was in fact equal. This was done to ensure that participants remained motivated to attend to the cues and outcomes.

After the scanning session, participants rated how positive they felt when observing themselves or the other person winning on a nine-point scale ranging from "not at all" to "very positive." One-sample t tests showed that, for both self and other, participants felt significantly more positive than neutral when seeing win outcomes compared with no win outcomes (other win $t_{(29)} = 2.1$, $p < 0.05$, M = 5.4, SD = 1.04; self win $t_{(29)} = 5.3$ $p < 0.001$, M = 6.4, SD = 1.43), and paired-sample t tests showed that participants felt significantly more positive when they won money for themselves compared with seeing the other participant win ($t_{(29)} = 4.35$, $p < 0.001$). This suggests that participants found it rewarding to view win outcomes for both themselves and for the other participant, but felt more positive overall when they viewed themselves winning.

Questionnaire measures. Participants completed a measure of empathy, the Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011). The QCAE is a multidimensional instrument devised to measure five key components of empathy. In the development of the QCAE, two raters selected items from other well validated and commonly used empathy measures (e.g., Empathy Quotient; Hogan Empathy Scale; the Empathy subscale of the Impulsiveness-Venturesomeness-Empathy Inventory; and the Interpersonal Reactivity Index) if they were deemed to measure empathy (see items below). Items deemed to measure other processes (e.g., sympathy) were not included. The five subscales comprising the QCAE are as follows: perspective-taking (e.g., "I can easily tell if someone else wants to enter a conversation."); online simulation (e.g., "Before criticizing somebody, I try to imagine how I would feel if I was in their place."); emotion contagion (e.g., "I am happy when I am with a cheerful group and sad when the others are glum."); peripheral responsivity (e.g., "I often get deeply involved with the feelings of a character in a film, play, or novel."); and proximal responsivity (e.g., "I often get emotionally involved with my friends' problems"). Items are rated on a four-point scale from "strongly disagree" to "strongly agree." The QCAE has good construct validity and internal consistency (Reniers et al., 2011).

Statistical analysis of behavioral data. Behavioral analyses were performed in SPSS 22 software (IBM). An agency (self vs other) by reward (win vs no win) ANOVA was used to examine reaction time (RT) differences to outcome judgments. We did not examine the agency (self vs other) by expectedness (expected vs unexpected) interaction due to the low number of unexpected outcomes in our design (<10 valid trials per subject). Relationships between behavioral performance and empathy were assessed using bivariate correlations. We adopted an α level of 0.05 and a power analysis indicated that we had ~80% power to detect an effect size of Cohen's $d = 0.50$.

Functional neuroimaging data collection and analysis

fMRI data acquisition. A Siemens Avanto 1.5-T MRI scanner was used to acquire a 5.5 min 3D T1-weighted structural scan and 424 multislice T2*-weighted echoplanar volumes with BOLD contrast. The structural scan was acquired using a magnetization prepared rapid gradient echo sequence. Imaging parameters were as follows: 176 slices; slice thickness = 1 mm; gap between slices = 0.5 mm; TR = 2730 ms; TE = 3.57 ms; field of view = 256 mm × 256 mm^2; matrix size = 256 × 256; voxel size = 1 × 1 × 1 mm resolution. The echoplanar image (EPI) sequence was acquired in an ascending manner, at an oblique angle (≈30°) to the AC–PC line to decrease the impact of susceptibility artifact in the orbitofrontal cortex (Deichmann et al., 2003) with the following acquisition parameters: 424 T2*-weighted echoplanar volumes, 35 2 mm slices, 1 mm slice gap; echo time = 50 ms; repetition time = 2975 ms; flip angle = 90°; field of view = 192 mm; matrix size = 64 × 64.

fMRI data analysis. Imaging data were analyzed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Data preprocessing followed a standard sequence. The first four volumes were discarded to allow for T1 equilibration effects and the last volume was discarded because the experimental task ended one volume before the end of the scanning sequence. The removal of the last volume ensured that no hemodynamic response (which typically occurs 4–6 s after event onset) to the desktop screen was sampled. Images were then realigned and coregistered to the participant's own anatomical image. The anatomical image was processed using a unified segmentation procedure combining segmentation, bias correction, and spatial normalization to the Montreal Neurological Institute (MNI) template using SPM's New Segment procedure (Ashburner and Friston, 2005); the same normalization parameters were then used to normalize the EPI images. The images were resampled to a voxel size of 1.5 × 1.5 × 1.5 mm. Finally, a Gaussian kernel of 8 mm full-width at half-maximum was applied to spatially smooth the images. Before the study, first-level design matrices were examined to ensure that estimable GLMs could be performed with independence between all regressors with correlation coefficients of $r < 0.25$.

First-level analysis. Nine (in some subjects, 10) event types were used to construct regressors in which event onsets were convolved with the synthetic canonical hemodynamic response function in SPM and associated responses were estimated in the context of the general linear model. Each of the four conditions (self high probability, self low probability, other high probability, other low probability) at the time of the cue and at the time of the outcome was modeled as a separate regressor for correct responses. The onset of the judgment was also modeled in a single regressor across all event types. An additional regressor modeled trials in which the judgment was missed or participants made an error. For those participants whose head motion caused visible image corruption in particular scans, an extra regressor was included. These images were removed and replaced with an image created by interpolating the two adjacent images to prevent distortion of the between-subjects mask (four participants, each accounting for <1% of the total fMRI data). The residual effects of head motion were also modeled as covariates of no interest in the analysis by including the six head motion parameters estimated during realignment. Data were high-pass filtered at 128 s to remove low-frequency drifts and the statistical model included an AR(1) autoregressive function to account for autocorrelations intrinsic to the fMRI time series. Contrast images were computed to examine the interaction (agency × probability) and main effects of agency (self > other and other > self) and probability (high > low and low > high) at the time of the cue.

Many studies have suggested that situations that involve mixed payoffs between study participants and other people can result in neural responses that reflect payoff differences between self and other; that is, they relate to coding of rewards for self relative for other, often called "inequity aversion" rather than "vicarious" reward responses (for reviews, see Ruff and Fehr, 2014; Rilling and Sanfey, 2011). To determine whether identified neural responses to reward predicting cues in the current study were reflective of coding of rewards for self relative to other, and thus inequity aversion, we constructed a second model that was the same as the main model but contained all cues collapsed into a single regressor. This regressor had two associated parametric modulators. The first coded the "inequity," the difference in accumulated reward between self and other on each trial, and the second coded the agent × probability interaction. This allowed us to determine neural responses to inequity and whether any neural responses occurred over and above the variance explained by inequity.

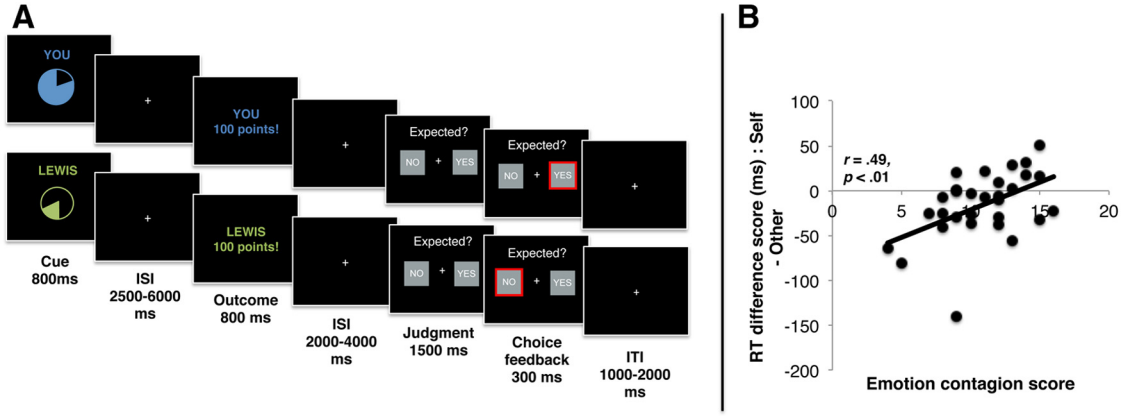


Figure 1. ***A***, Trial structure. Participants performed trials that began with a cue signaling the probability of reward (high [80%] or low [20%]) and the agent to whom reward would be delivered (self = "you" and other = "Lewis" in this example). Participants judged whether the outcome (win 100 points or win 0 points) was expected or unexpected after outcome delivery. Participants believed that the other participant outside of the scanner was simultaneously performing the same task and that the points they observed would be converted into additional payment at the end of the experiment for themselves and for the other participant. ***B***, Scatterplot showing association between self–other RT difference at the time of the judgment and trait emotion contagion ($n = 30$). Overall, participants were slower when making judgments about the expectedness of outcomes for other compared with self. However, this effect was associated with emotion contagion such that those highest in emotion contagion showed a relative speeding of response for other.

Second-level analysis. Second-level analysis was performed using the standard summary statistics approach to random-effects analysis in SPM. Contrast images were input into second-level one-sample *t* test design matrices. Interactions and main effects are reported at $p < 0.05$, familywise error (FWE) corrected at the voxel level across the whole brain. Where significant interactions were identified, we conducted illustrative *post hoc* analyses with simple main effects contrasts using a less conservative statistical threshold of $p < 0.001$ (uncorrected).

Results

Behavioral data

Participants were highly accurate in their judgments of whether the outcome was expected or not (mean accuracy >91% for all participants for both trial types) and missed very few trials (mean <1% for all participants). For mean RTs, a 2 (self vs other) by 2 (win vs no win) ANOVA showed significantly slower judgments on third-person (M = 664 ms, SD = 18) than on first-person (M = 649 ms, SD = 16) trials (main effect of agency: $F_{(1,29)} = 5.32$, $p = 0.03$). Judgments were also significantly faster after reward (641 ms, SD = 16) compared with a no reward (672 ms, SD = 19) outcomes (main effect of outcome $F_{(1,29)} = 14.34$, $p = 0.001$). The agency × reward interaction was nonsignificant ($F_{(1,29)} = .05$, $p = 0.83$).

Given the significant main effect of agency, we calculated the difference score between self and other RTs to examine associations between this behavioral measure and empathy. The emotion contagion subscale of the QCAE was positively associated with the self–other difference score ($r = 0.49$, $p < 0.01$); that is, participants higher in emotion contagion showed a relative facilitation (speeding) when making decisions about the expectedness of outcomes for other people (Fig. 1*B*). No other subscale of the QCAE correlated with the self-other difference score (all $p > 0.49$). Multiple regression, including all QCAE subscales, showed that the association between the self–other difference score and self-reported empathy was specific to the emotion contagion subscale ($\beta = 0.55$, SEM = 2.43, $p < 0.01$).

fMRI data

Agency × probability interaction at time of the cue

To test our first hypothesis, that activity in the ACCg would signal information about reward probability for others, we examined the agency × probability interaction at the time of the cue. Consistent with our hypothesis, this analysis revealed a significant effect in the ACCg (MNI coordinates [$x = 8$, $y = 32$, $z = 12$], $Z = 5.05$, $k = 10$, $p < 0.05$ FWE, whole brain corrected), putatively in area 24a′/24b′ at the border of the midcingulate and anterior cingulate subregions (Fig. 2). We examined the nature of this interaction by testing the simple main effects, specifically the contrasts of other high versus low probability and self low versus high probability. Inspection of the other high versus low probability simple main effect revealed a large cluster in the ACCg overlapping with the region identified in the interaction (MNI coordinates [$x = 6$, $y = 33$, $z = 12$], $Z = 4.14$, $k = 184$, $p < 0.001$, uncorrected). Inspection of the self low versus high probability contrast revealed a small cluster of overlapping voxels (MNI coordinates [$x = 9$, $y = 32$, $z = 13$], $Z = 3.28$, $k = 5$, $p < 0.001$, uncorrected). This suggests that the ACCg activation identified in the interaction mainly signals the probability of rewards that would be received by another person.

Associations with trait empathy

To test our second hypothesis, that the extent to which ACCg responds to the probability of rewards specifically for others would be positively associated with trait empathy, we used MarsBaR (Brett et al., 2002) to extract individual interaction contrast estimates (other high vs low probability minus self high vs low probability) from the ACCg cluster identified above and correlated these with participants' self-reported empathy on the five QCAE subscales. Emotion contagion was significantly negatively associated with the ACCg interaction contrast estimate ($r = -0.45$, $p = 0.01$, all other subscales $p > 0.58$) and multiple regression, including all QCAE subscales, showed that this effect was specific to emotion contagion ($\beta = -.60$, SEM = 0.062, $p = 0.003$, all other subscales $p > 0.15$; Fig. 3). In other words, the interaction was weakest in individuals high in emotion contagion.

To better understand the nature of this association, we examined the correlations for other high versus low probability and self low versus high probability in ACCg with empathy subscales (Fig. 3). There was no significant correlation between ACCg response to other high versus low probability ($r = -0.05$, $p = 0.81$) and empathy. However, there was a significant negative association between ACCg response to self low versus high probability and emotion contagion ($r = -0.58$, $p < 0.001$); again, multiple

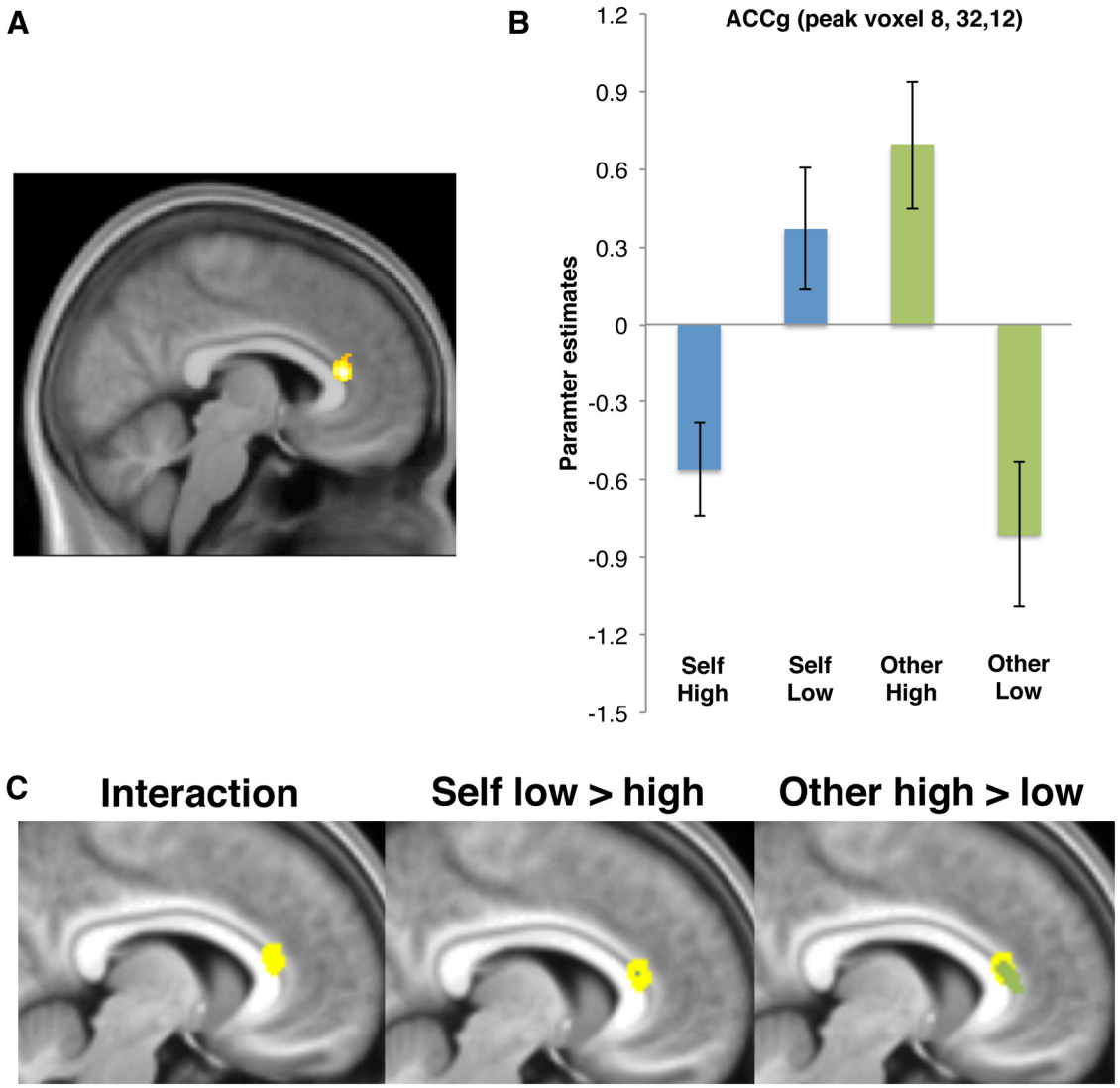


Figure 2. ***A***, Activation in the ACCg signaled the agency (self vs other) by probability (high [80%] or low [20%]) interaction at the time of the cue [$x = 8, y = 32, z = 12$], displayed at $p < 0.001$ (uncorrected). ***B***, Parameter estimates for the peak voxel in the ACCg. ***C***, Left, Overlay of the agency × probability interaction in ACCg (yellow, as in ***A***). Middle, Only a small number of voxels overlapped between the interaction contrast (yellow) and the simple main effect of self low versus high probability (blue, $k = 5$ at $p < 0.001$ uncorrected). Right, A large number of voxels overlapped between the interaction contrast (yellow) and the simple main effect of other high > low probability (green, $k = 184$ at $p < 0.001$ uncorrected). Error bars indicate SEM.

regression demonstrated that this effect was unique to emotion contagion ($\beta = -0.66$ SEM $= 0.082$, $p < 0.001$, all other subscales $p > 0.19$). In other words, the extent to which ACCg distinguished between low and high reward probability for self was attenuated in individuals with high emotion contagion.

In summary, in individuals with high emotion contagion, the ACCg signaled information about the relative difference between high and low probability rewards only for others, whereas in individuals with low emotion contagion, the ACCg additionally signaled (negatively) reward probability for self.

Main effects at the time of the cue

The temporal pole showed a significant main effect of other > self (MNI coordinates [33, 22, −26]; $Z = 4.85$; $k = 2$, $p < 0.05$, FWE whole brain corrected). No other main effects or interactions survived whole-brain correction for multiple comparisons.

Agency × outcome interaction and main effects at the time of the outcome

No interactions or main effects survived whole-brain correction for multiple comparisons.

Analysis of inequity aversion

Analysis of the inequity parametric modulator showed no whole-brain-corrected results and no uncorrected results in ACCg. We then tested whether our observed effects in the ACCg occurred over and above any effects of inequity. This analysis showed that there was still a significant effect in the ACCg after accounting for the variance explained by inequity (MNI coordinates [$x = 6, y = 32, z = 13$], $Z = 4.97$, $k = 8$, $p < 0.05$ FWE, whole brain corrected). Therefore, the ACCg response was unlikely to reflect differences in accumulated reward between self and other.

Discussion

We examined hemodynamic responses in the human brain to cues that predicted a high or low probability of a reward for oneself or another person. We show that the ACCg robustly signals the probability of rewards for another person. This supports our hypothesis that the ACCg is engaged when processing predictions about rewards for other people. Our second hypothesis, that that the extent to which the ACCg is specialized for processing others' rewards is positively associated with trait empathy,

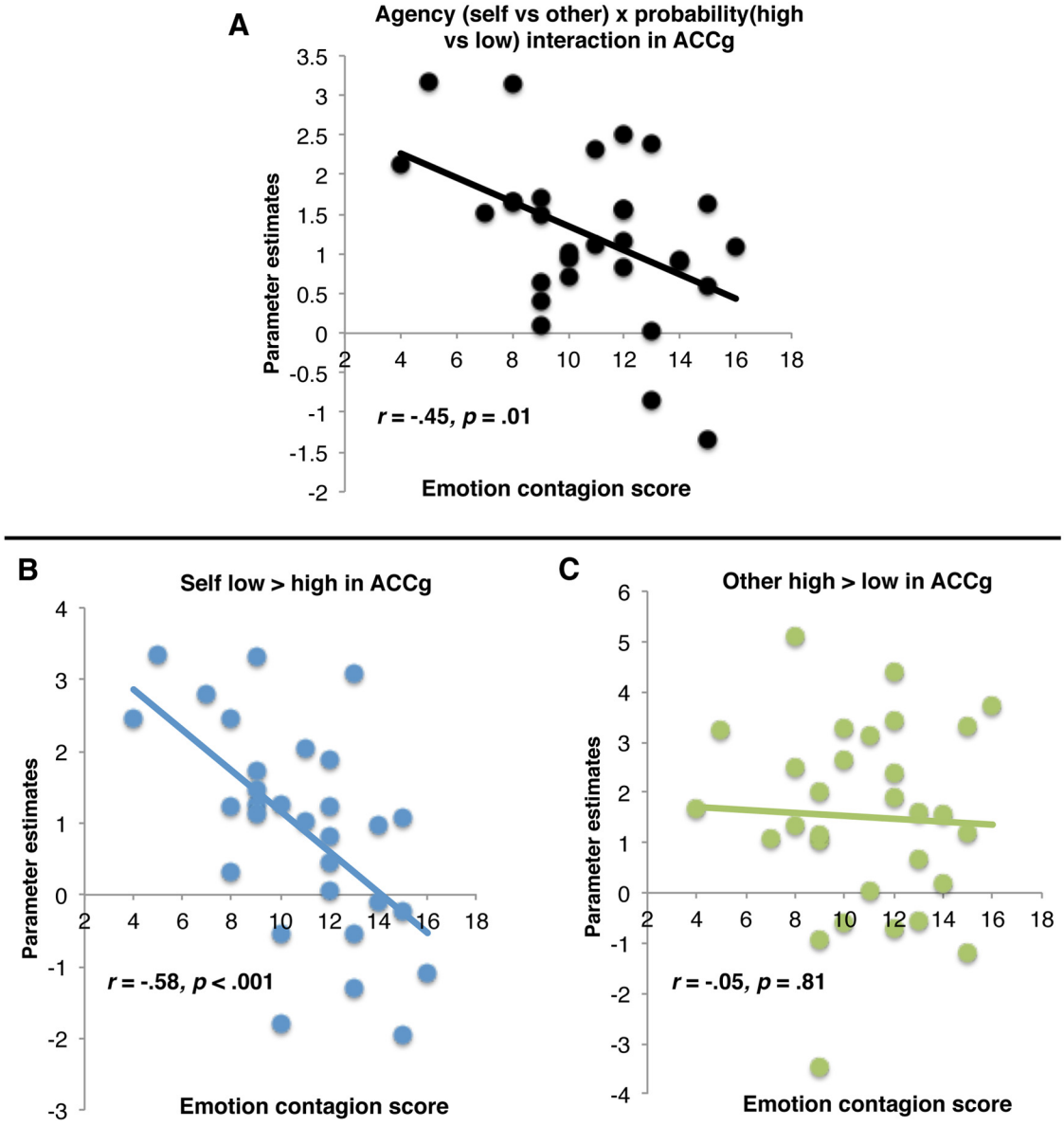


Figure 3. ***A***, Significant association between the cluster in the ACCg showing the interaction effect and participants' emotion contagion scores. ***B***, Response to self low > high probability decreases as a function of emotion contagion, with those lowest in emotion contagion showing the greatest response to low > high probability of reward for self. ***C***, Response to other high > low probability shows no significant modulation as a function of emotion contagion.

was partially supported. As predicted, the interaction effect in the ACCg significantly covaried with emotion contagion. However, this effect was driven by the extent to which ACCg signaled reward predictions for self, not other. Specifically, for those high in emotion contagion, the ACCg signaled reward prediction exclusively for others, whereas for those low in emotion contagion, this same region signaled reward prediction for self (in the opposite direction).

The model of the contributions of ACCg to social cognition (Apps et al., 2013b) highlights that this region plays an important role in understanding the value of others' rewards, and consequently in social behavior (Rudebeck et al., 2006; Behrens et al., 2008; Jones et al., 2011; Apps et al., 2013a, 2015; Boorman et al., 2013; Chang et al., 2013; Apps and Ramnani, 2014). This claim is built upon several lines of evidence. Lesions to this region have been shown to impair the processing of social stimuli and cause a reduction in the execution of social behaviors (Rudebeck et al., 2006). The ACCg is connected to regions that process social information, but also to regions that process reward-related information (Yeterian and Pandya, 1991; Lynd-Balta and Haber, 1994; Haber et al., 1995). Single-unit recording evidence suggests that a relatively large proportion of ACCg neurons, compared with those in other prefrontal regions, respond when a monkey anticipates the delivery of reward to another monkey (Chang et al., 2013), and human imaging studies have shown that the ACCg responds when tracking the value of cues predicting approval from peers (Jones et al., 2011). Together, these studies support the claim that the ACCg is important for processing others' rewards and also in social behavior. However, a key untested component of this model was that the ACCg would be engaged when processing the likelihood of rewards being delivered to others. We show for the first time that the ACCg signals the likelihood of others' rewards regardless of trait levels of empathy. We also note that we did not observe responses to reward prediction in other

candidate regions for reward signals, even at uncorrected levels (e.g., ventral striatum, ventromedial prefrontal cortex, and amygdala; for a meta-analysis, see Morelli et al., 2015), supporting some degree of specificity of ACCg response to vicarious reward in our study.

Our experimental paradigm was designed to ensure that participants attended to reward cues. By asking participants to make a decision at the time of the outcome, we cannot purely assess whether outcome-related responses are also coded in ACCg because participants were both processing the outcome and preparing a motor response during this time. However, there is evidence that vicarious prediction error signals may well be coded in ACCg (Apps et al., 2013a, 2015). We provide the first evidence that this same region also encodes the likelihood of others receiving rewards.

Although previous studies have suggested the ACCg plays an important role in empathy (Lamm et al., 2011; Engen and Singer, 2013), these studies have largely focused on neural responses to others' pain. Our data suggest that the degree of specialization in this region's response to others' predicted rewards may partly underlie individual differences in emotion contagion. Emotion contagion is hypothesized to be a necessary foundation for empathizing with other individuals (Bird and Viding, 2014) and is a process that is shared with nonhuman animals (for review, see de Waal, 2008). Importantly, emotion contagion also covaried with RTs to decisions about rewards delivered to others, with those highest in trait emotion contagion showing the greatest speeding of response. A distinction is often made between "affective empathy," which is commonly understood as an affective state caused by vicariously processing the experiences of another person, and "cognitive empathy," which is thought to include processes such as perspective taking and theory of mind (Singer and Lamm, 2009). Regression analyses suggested that only emotion contagion, part of the "affective" component, was associated with vicarious reward prediction. In tasks investigating cognitive aspects of empathy, an anatomically separate region of the mPFC, the dorsal mPFC, is often responsive (Amodio and Frith, 2006), suggesting partially separate functions of the ACCg and mPFC.

Although we did not predict an association between emotion contagion and ACCg response to self reward prediction, a possible explanation relates to the findings of Chang et al. (2013) and Haroush and Williams (2015). These investigators observed some self-reward- selective neurons in the same region of the ACCg/dACC that also contained other-reward-selective neurons, suggesting that some processing of information about rewards for self occurs in ACCg. However, given the limited sample sizes in nonhuman primate studies, the investigators were unable to examine variability in the proportion of neurons that signaled self versus other reward. We speculate that, even if at the population level, the ACCg shows a relative specialization in processing rewards for others, individual variability in the degree to which self rewards are also processed in this region could be important for explaining heterogeneity in ACCg function and empathy. That is, for those individuals who display the lowest levels of emotion contagion, there appears to be reduced specialization and a potentially opposing coding scheme of self and other reward probability in ACCg. Such opposing coding within the same anatomical region could have consequences for understanding social cognition and behavior, such as increased weighting of rewards to self and higher likelihood of engaging in competitive social interactions.

This interpretation is supported by a recent study finding that stimulation of dACC neurons made monkeys more competitive (Haroush and Williams, 2015). Similarly, another study showed that single neurons in a region of the rat cingulate cortex thought to be homologous with human dACC coded the value of competing with another rat for rewards (Hillman and Bilkey, 2012). These findings may help to reconcile previous discrepancies in the functions that have been imputed to dACC in terms of competitive social behaviors (Hillman and Bilkey, 2012; Haroush and Williams, 2015), but also empathy (Lamm et al., 2011; Engen and Singer, 2013). We propose that variability in empathy may modulate, not only the extent to which social information is processed in ACCg, as suggested in previous studies and theoretical accounts of empathy (Lamm et al., 2011; Engen and Singer, 2013), but also the extent to which self and other reward information is computed. However, this hypothesis requires further testing in future experiments.

Empathic abilities are a fundamental building block for successful social behavior and are at the core of many disorders of social cognition, including autism and psychopathy (Blair, 2005; Lockwood et al., 2013a; Bird and Viding, 2014). Previous studies have suggested that a similar portion of the dACC that was activated in our study is anatomically and functionally atypical in individuals with psychopathy and in individuals with autism (Simms et al., 2009; Brazil et al., 2011; Anderson and Kiehl, 2012; Delmonte et al., 2013; Lockwood et al., 2013b). Integrating these previous findings with the present results suggests the hypothesis that individual differences in the structure, function, and connectivity of the ACCg constrain the extent to which this region processes reward-predicting cues for others compared with self, which may lead to atypical empathic processing. However, we also know that psychopathy and autism have different profiles of empathic processing and behavior from one another (Blair, 2005; Lockwood et al., 2013a; Bird and Viding, 2014). The ACCg has strong connections to other regions involved in social and reward processing, including the nucleus accumbens (Yeterian and Pandya, 1991; Lynd-Balta and Haber, 1994; Haber et al., 1995), a region also suggested to participate in vicarious reward processing (Mobbs et al., 2009; Fareri et al., 2012; Braams et al., 2014), the temporal poles (which showed greater response to other vs self reward prediction in our study),and the temporoparietal junction and paracingulate cortex (Markowitsch et al., 1985; Seltzer and Pandya, 1989; Barbas et al., 1999). Future research into the neurocognitive correlates of psychopathy and autism should investigate whether distinct social behavioral abnormalities can be characterized by differences in the functional and connectional fingerprint of the ACCg during vicarious reward processing.

In summary, we demonstrate a central role for the ACCg in processing predictions about the likelihood of others' rewards. We also found substantial individual variation in the degree to which the ACCg responds to self and other reward, with only those highest in trait emotion contagion showing specialization of ACCg for others predicted reward. Together, our findings highlight the importance of understanding the contributions of the ACCg to social cognition and how variability in its function may underlie variability in social behavior.

Notes

Supplemental material for this article is available at http://www.patricialockwood.co.uk/Publications. In Table 1, we provide uncorrected results at the time of the cue ($p < 0.001$, $k = 10$) for completeness. We note that these results should be interpreted with caution given that they do not survive correction for multiple comparisons. In Table 2, we

provide uncorrected results at the time of the outcome ($p < 0.001$, $k = 10$) for completeness. We note that these results should be interpreted with caution given that they do not survive correction for multiple comparisons. This material has not been peer reviewed.

References

Amodio DM, Frith CD (2006) Meeting of minds: the medial frontal cortex and social cognition. Nat Rev Neurosci 7:268–277. CrossRef Medline

Anderson NE, Kiehl KA (2012) The psychopath magnetized: insights from brain imaging. Trends Cogn Sci 16:52–60. CrossRef Medline

Apps MA, Ramnani N (2014) The anterior cingulate gyrus signals the net value of others' rewards. J Neurosci 34:6190–6200. CrossRef Medline

Apps MA, Green R, Ramnani N (2013a) Reinforcement learning signals in the anterior cingulate cortex code for others' false beliefs. Neuroimage 64:1–9. CrossRef Medline

Apps MA, Lockwood PL, Balsters JH (2013b) The role of the midcingulate cortex in monitoring others' decisions. Front Neurosci 7:251. Medline

Apps MA, Lesage E, Ramnani N (2015) Vicarious reinforcement learning signals when instructing others. J Neurosci 35:2904–2913. CrossRef Medline

Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839–851. CrossRef Medline

Barbas H, Ghashghaei H, Dombrowski SM, Rempel-Clower NL (1999) Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. J Comp Neurol 410:343–367. CrossRef Medline

Behrens TE, Hunt LT, Woolrich MW, Rushworth MF (2008) Associative learning of social value. Nature 456:245–249. CrossRef Medline

Bird G, Viding E (2014) The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. Neurosci Biobehav Rev 47:520–532. CrossRef Medline

Blair RJ (2005) Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. Conscious Cogn 14:698–718. CrossRef Medline

Boorman ED, O'Doherty JP, Adolphs R, Rangel A (2013) The behavioral and neural mechanisms underlying the tracking of expertise. Neuron 80:1558–1571. CrossRef Medline

Braams BR, Güroğlu B, de Water E, Meuwese R, Koolschijn PC, Peper JS, Crone EA (2014) Reward-related neural responses are dependent on the beneficiary. Soc Cogn Affect Neurosci 9:1030–1037. CrossRef Medline

Brazil IA, Mars RB, Bulten BH, Buitelaar JK, Verkes RJ, De Bruijn ER (2011) A neurophysiological dissociation between monitoring one's own and others' actions in psychopathy. Biol Psychiatry 69:693–699. CrossRef Medline

Brett M, Anton J-L, Valabregue R, Poline J-B (2002) Region of interest analysis using the MarsBar toolbox for SPM 99. Neuroimage 16:S497.

Chang SW, Gariépy JF, Platt ML (2013) Neuronal reference frames for social decisions in primate frontal cortex. Nat Neurosci 16:243–250. Medline

de Waal FB (2008) Putting the altruism back into altruism: the evolution of empathy. Annu Rev Psychol 59:279–300. CrossRef Medline

Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. Neuroimage 19:430–441. CrossRef Medline

Delmonte S, Gallagher L, O'Hanlon E, McGrath J, Balsters JH (2013) Functional and structural connectivity of frontostriatal circuitry in autism spectrum disorder. Front Hum Neurosci 430.

Engen HG, Singer T (2013) Empathy circuits. Curr Opin Neurobiol 23:275–282. CrossRef Medline

Fareri DS, Niznikiewicz MA, Lee VK, Delgado MR (2012) Social network modulation of reward-related signals. J Neurosci 32:9045–9052. CrossRef Medline

Gabay AS, Radua J, Kempton MJ, Mehta MA (2014) The Ultimatum Game and the brain: A meta-analysis of neuroimaging studies. Neurosci Biobehav Rev 47:549–558. CrossRef Medline

Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. J Neurosci 15:4851–4867. Medline

Haroush K, Williams ZM (2015) Neuronal prediction of opponent's behavior during cooperative social interchange in primates. Cell:1–13.

Hillman KL, Bilkey DK (2012) Neural encoding of competitive effort in the anterior cingulate cortex. Nat Neurosci 15:1290–1297. CrossRef Medline

Jones RM, Somerville LH, Li J, Ruberry EJ, Libby V, Glover G, Voss HU, Ballon DJ, Casey BJ (2011) Behavioral and neural properties of social reinforcement learning. J Neurosci 31:13039–13045. CrossRef Medline

Lamm C, Decety J, Singer T (2011) Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. Neuroimage 54:2492–2502. CrossRef Medline

Lockwood PL, Bird G, Bridge M, Viding E (2013a) Dissecting empathy: high levels of psychopathic and autistic traits are characterised by difficulties in different social information processing domains Front Hum Neurosci 7:760.

Lockwood PL, Sebastian CL, McCrory EJ, Hyde ZH, Gu X, De Brito SA, Viding E (2013b) Association of callous traits with reduced neural response to others' pain in children with conduct problems. Curr Biol 23:901–905. CrossRef Medline

Lynd-Balta E, Haber SN (1994) Primate striatonigral projections–a comparison of the sensorimotor-related striatum and the ventral striatum. J Comp Neurol 345:562–578. CrossRef Medline

Markowitsch HJ, Emmans D, Irle E, Streicher M, Preilowski B (1985) Cortical and subcortical afferent connections of the primates temporal pole–a study of rhesus-monkeys, squirrel-monkeys, and marmosets. J Comp Neurol 242:425–458. CrossRef Medline

Mobbs D, Yu R, Meyer M, Passamonti L, Seymour B, Calder AJ, Schweizer S, Frith CD, Dalgleish T (2009) A key role for similarity in vicarious reward. Science 324:900. CrossRef Medline

Morelli SA, Sacchet MD, Zaki J (2015) Common and distinct neural correlates of personal and vicarious reward: a quantitative meta-analysis. Neuroimage 112:244–253. CrossRef Medline

Reniers RL, Corcoran R, Drake R, Shryane NM, Völlm BA (2011) The QCAE: a questionnaire of cognitive and affective empathy. J Pers Assess 93:84–95. CrossRef Medline

Rilling JK, Sanfey AG (2011) The neuroscience of social decision-making. Annu Rev Psychol 62:23–48. CrossRef Medline

Rogers RD, Ramnani N, Mackay C, Wilson JL, Jezzard P, Carter CS, Smith SM (2004) Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. Biol Psychiatry 55:594–602. CrossRef Medline

Rudebeck PH, Buckley MJ, Walton ME, Rushworth MF (2006) A role for the macaque anterior cingulate gyrus in social valuation. Science 313:1310–1312. CrossRef Medline

Ruff CC, Fehr E (2014) The neurobiology of rewards and values in social decision making. Nat Rev Neurosci 15:549–562. CrossRef Medline

Sallet J, Quilodran R, Rothé M, Vezoli J, Joseph JP, Procyk E (2007) Expectations, gains, and losses in the anterior cingulate cortex. Cogn Affect Behav Neurosci 7:327–336. CrossRef Medline

Schultz W (2013) Updating dopamine reward signals. Curr Opin Neurobiol 23:229–238. CrossRef Medline

Seltzer B, Pandya DN (1989) Frontal-lobe connections of the superior temporal sulcus in the rhesus-monkey. J Comp Neurol 281:97–113. CrossRef Medline

Shidara M, Richmond BJ (2002) Anterior cingulate: Single neuronal signals related to degree of reward expectancy. Science 296:1709–1711. CrossRef Medline

Simms ML, Kemper TL, Timbie CM, Bauman ML, Blatt GJ (2009) The anterior cingulate cortex in autism: heterogeneity of qualitative and quantitative cytoarchitectonic features suggests possible subgroups. Acta Neuropathol 118:673–684. CrossRef Medline

Singer T, Lamm C (2009) The social neuroscience of empathy. Ann N Y Acad Sci 1156:81–96. CrossRef Medline

Vogt BA, Nimchinsky EA, Vogt LJ, Hof PR (1995) Human cingulate cortex–surface-features, flat maps and cytoarchitecture. J Comp Neurol 359:490–506. CrossRef Medline

Yeterian EH, Pandya DN (1991) Prefrontostriatal connections in relation to cortical architectonic organization in rhesus-monkeys. J Comp Neurol 312:43–67. CrossRef Medline