Motive, Character and the Moral Good:  
Elaboration on a Footnote

T.J. Philip

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Otago, Dunedin

New Zealand

2014
Abstract

Research on moral judgement has mostly examined judgements of moral wrongdoing. Little has been said about how people decide if an action is morally good. Furthermore there is evidence that judgements of good and bad actions differ sufficiently to make generalizing from the literature on moral wrongdoing to the moral good, problematic. This thesis examines the concept of a morally good action, focussing on how an agent’s motive and control over their actions influence an observer’s goodness judgements. Study 1 takes a qualitative look at the concept of moral goodness and identifies a number of factors central to the folk concept of a morally good action, including the importance of an agent’s motive. Study 2 provides experimental evidence that the motives identified in Study 1 (the interest of others, principles and self-interest) influence goodness judgements of actions, and that this relationship is partially mediated by an observer’s perception of an agent’s character. Studies 3-5 examine how good actions resulting from sympathy influence an observer’s goodness judgments. Finally Studies 6-8 examine how an agent’s control over their actions influences an observer’s goodness judgements, and find that loss of control from an internal source reduces how morally good an action is judged to be. This thesis is an exploratory look at the concept of a morally good action, and through a series of studies advances understanding of the relation between motive, character, control, and goodness judgements of actions.
Acknowledgments

My thanks to Jamin Halberstadt for all he has taught me over the years. Thanks to Tamlin Conner for her help with editing, Rachel Butler and Luxi Nie for their coding expertise, and the members of the Social Cognition Lab, in particular Jonathan Jong and Matthias Blümke for their thoughtful discussions and friendship. Thank you to my parents for their support and to my mother, Robyn Philip, for all the editing. Thanks Mike and Karen Knapp for taking me in when I was homeless and Simon and Rebecca Walker for looking out for me. I also want to thank Kristen Steslow for nudging my wretch-husk of a corpse in the direction of humanity.
Table of Contents

Abstract ............................................................................................................................ i
Acknowledgments ......................................................................................................... ii
Table of Contents .......................................................................................................... iii
List of Tables & Figures ................................................................................................. vi

Chapter 1 ....................................................................................................................... 1
Two kinds of morality .................................................................................................... 1
  Psychological theories of moral judgement ............................................................... 3
  What do psychologists think about the moral good? ................................................. 5
  Transgressions and dilemmas ................................................................................... 7
  Two kinds of judgements ......................................................................................... 8
  Conclusion ................................................................................................................ 10

Chapter 2 ....................................................................................................................... 11
The folk concept of moral goodness .......................................................................... 11
  Folk concepts .......................................................................................................... 11
  The folk concept of moral goodness .................................................................... 12
  Study 1 .................................................................................................................... 13
    Method .................................................................................................................. 13
    Results and Discussion ....................................................................................... 14
    Motives ................................................................................................................ 20

Chapter 3 ....................................................................................................................... 27
Motives and goodness judgements .......................................................................... 27
  Mind is central to morality ..................................................................................... 27
  Mind and behaviour explanations ......................................................................... 28
  Behaviour explanations and moral judgement .................................................... 29
  Intention and control ............................................................................................. 29
  Motives .................................................................................................................... 30
  Motives as indication of character? ..................................................................... 34
  Study 2 .................................................................................................................... 35
    Method .................................................................................................................. 36
    Results .................................................................................................................. 38
    Discussion ............................................................................................................ 42
    Conclusion ............................................................................................................ 46

Chapter 4 ....................................................................................................................... 48
Sympathy and moral goodness .................................................................................... 48
  Sympathy ................................................................................................................ 48
  Study 3 .................................................................................................................... 50
    Method .................................................................................................................. 50
    Results .................................................................................................................. 52
    Discussion ............................................................................................................ 54
  Study 4a ................................................................................................................. 56
    Method .................................................................................................................. 57
    Results .................................................................................................................. 57
    Discussion ............................................................................................................ 58
List of Tables & Figures

Table 2. 1. Categories derived from participant’s responses to what is morality........... 15
Table 2. 2. Category frequency and coder agreement................................................. 18
Table 3.1. Example vignettes for each condition ...................................................... 38
Table 3.2. Path coefficients and significance for each pair wise comparison ............ 42
Table 4.1. Example vignettes for volunteered and questioned vignette types, and sympathy and no sympathy conditions.......................................................... 51
Table 4.2. Mediation analysis between sympathy condition and goodness judgements by character judgement................................................................. 60
Table 4.3. Example volunteered vignettes for sympathy and no sympathy conditions. 62
Table 4.4. Mediation analysis between sympathy condition and goodness judgements by character judgements ........................................................................ 63
Table 4.5. Example questioned vignettes for sympathy and no sympathy conditions ... 65
Table 4.6. Mediation analysis between sympathy condition and goodness judgements by character judgements ........................................................................ 66
Table 5.1. Sample vignettes from Pizarro, Uhlmann & Salovey (2003)....................... 72
Table 5.2. Example vignettes for frontal lobe damage manipulation.............................. 77
Table 5.3. Example vignettes for impulse control difficulty manipulation.................. 78
Table 5.4. Questionnaire versions ............................................................................... 80
Table 5.5. Full design to test the asymmetry hypothesis.............................................. 88
Table 5.6. Example stimuli for control manipulations.................................................. 90

Figure 3.1. Mean moral judgements and character scores. Scales range from 1 to 9 with 5 as the centre point................................................................. 40
Figure 3.2. Model 1. Overall effect of an agent’s motive on an observer’s goodness judgements................................................................. 41
Figure 4.1. Mean moral goodness scores or sympathy and no sympathy condition for
volunteered and questioned vignettes................................................................. 54

*Figure 5.1.* Frontal lobe manipulation. Mean moral judgement score for morally good and bad acts that were either deliberate or due to an impulse. .......................... 79

*Figure 5.2.* Impulse control history manipulation. Mean moral judgement score for morally good and bad acts that were either deliberate or due to an impulse......... 81

*Figure 5.3.* Study 7. Mean moral judgement scores for morally good and bad acts that were either deliberate or due to an impulse ...................................................... 85

*Figure 5.4.* Meta-desire scores for morally good and bad, deliberate and impulsive conditions ........................................................................................................ 86

*Figure 5.5.* Mean moral judgement scores for deliberate and anger conditions displayed by vignette and order. ................................................................................... 91
Chapter 1

Two kinds of morality

Morality is most commonly described in the psychological literature as standards of conduct to which people are required to adhere. Haidt (2007) describes a moral community as having “a set of shared norms about how members ought to behave” (p.1000). Similarly, De Wall (1996, p.10) writes that morality is “commonly cast as how we should or should not behave in order to be valued members of society”, and Janoff-Bulman, Sheikh, and Hepp (2009, p. 521) describe morality as “standards of conduct developed to coordinate and facilitate group living”. While standards of behaviour are certainly an important part of morality, there is evidence that morality also involves concerns for “rights and duties, good and bad character traits (virtues and vices), and right and wrong motives” (Krebs, 2008, p.150, see also Haidt, 2001; Inbar, Pizarro, & Cushman, 2012; Woolfolk, Doris, & Darley, 2006).

Moral judgement, is then a positive or negative evaluation with respect to certain shared standards. But theories of what these standards are, and how they can be distinguished from non-moral social standards vary greatly among researchers. Early researchers into morality – whose main focus was on moral development – considered morality to be concerned with issues of harm and justice (Kohlberg, 1969; Turiel, 1983, Nucci, 2001). Later research by Shweder (1990; Shweder, Much, Mahapatra, & Park, 1997) and subsequently Haidt and colleagues (Haidt & Graham, 2007; Haidt & Joseph, 2004) expanded the domain of morality beyond concerns for harm and fairness to include authority, loyalty and purity (but see Suhler & Churchland, 2011). In contrast to these narrower characterizations of morality, Rai and Fiske (2011) argue that moral judgements cannot be considered independently of the social context in which they occur and that researchers should abandon domain-specific descriptions of morality entirely (see also Parkinson et al., 2011).

Though theories differ on what exactly constitutes a moral judgement, researchers generally agree that qualitative differences exist between moral and other kinds of non-moral social evaluations. Research suggests that people treat moral beliefs more like objective facts about the world (Goodwin & Darley, 2008). Moral wrongness – and presumably rightness – is perceived to be less dependent on societal agreement or
the decree of an authority than ‘conventional’ non-moral transgressions (Huebner, Lee, & Hauser, 2010, Nucci, 2001), and moral judgements are considered to be more universal. That is, an action is morally wrong regardless of when or where it happened (Nucci, 2001; Sousa, Holbrook, & Piazza, 2009). Consequently, moral standards have prescriptive force, and adherence is obligatory (Haidt, 2001; Nucci, 2001; Shweder, Turiel, & Much, 1981; Sousa, Holbrook, & Piazza, 2009; but see Kelley, Stich, Haley, Eng & Fessler, 2007 & Stich, Fessler, & Kelly, 2009). It has also been demonstrated that moral beliefs resist being compromised or traded off for other non-moral interests (Baron & Spranca, 1997; Tetlock, Kristel, Elson, Greene, & Lerner, 2000), and cannot be solely predicted by the usual attitudinal dimensions, suggesting that moral attitudes are not just very strong attitudes (Skitka, Bauman, & Sargis, 2005), though there is evidence that moral judgements are associated with stronger emotions than social judgements (Skitka, 2010). Furthermore, there is evidence for motivational effect of moral beliefs; Wright, Cullum, and Schwab (2008) found that people are intolerant of and share less with people whose moral beliefs differed from theirs.

Moral judgements can be expressed in a number of ways. Saying something is good, right or ought to be done, are all ways to express moral judgement (Hare, 1963). Cushman (2008) uses the term moral judgement to “refer globally to a broad class of evaluations that include wrongness, permissibility, punishment, blame and many other specific types of judgment” (p.355) which is characteristic of the treatment these ideas receive in the literature on moral psychology. Though there is a particular body of literature identified as researching moral judgements, other literatures deal with related and overlapping issues. For instance, the blame literature examines factors relevant to ascription of blame, which often identifies moral wrongdoing (Alicke, 2000). The attribution literature is concerned with inferring traits from actions (Jones & Davis, 1965; Uleman, Newman, & Moskowitz, 1996), and evaluative traits are often moral evaluations (Narvaez, Lapsley, Hagele, & Lasky, 2006). Furthermore, there is evidence that the moral judgement and trait attribution literatures are more closely related than they may at first appear. Hare (1963) points out that “when we use the word ‘good’ in order to commend morally, we are always directly or indirectly commending people. Even when we use the expression ‘good act’ or others like it, the reference is indirectly to human characters” (p.144). Research suggests that perhaps moral judgements that
appear at first glance to be evaluations of actions also serve as evaluations of character (Inbar, Pizarro, & Cushman, 2012; Pizarro & Tannenbaum, 2011).

Psychological theories of moral judgement

Early theories of moral judgement focused on moral reasoning and justification (Kohlberg, 1969; Turiel, 1983). Kohlberg’s work identified stages of moral development; in the earliest stage, right and wrong are justified by the possibility of punishment, while the final stage is reached when people justify their moral judgements by appealing to abstract universal principles. There is ongoing work in this tradition, and as with the original research, the emphasis remains on how children reason about morality (Malti, Gasser, & Buchman, 2009; Nucci, 2001).

More recent research has questioned the role of reasoning in making moral judgements, and instead emphasizes the importance of emotion and automatic processes (Haidt, 2001). Championing this approach, Haidt argues for the primacy of emotion based intuition in moral judgement (2001, 2007). According to the social intuitionist model, moral judgements are predominantly made using non-conscious – primarily affective – processes, and that moral reasoning is mostly post hoc justification of moral intuitions. Supporting this argument a number of studies have found evidence for the importance of emotion in moral judgement (Haidt, 2003b; Inbar, Pizarro, Knobe, & Bloom, 2009; Russell & Giner-Sorolla, 2011; Schnall, Haidt, Clore, & Jordan, 2008; Schnall, Roper, & Fessler, 2010; Valdesolo & Desteno, 2006; Wheatley & Haidt, 2005).

Subsequently, dual process models have emerged that acknowledge the contribution of both conscious and non-conscious processes (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). For example Greene et al. (2001) use a dual process framework to explain the observation that people are more likely to judge a utilitarian decision as unacceptable (e.g., performing an action that will kill one person to save five) if the action performed involves direct physical or ‘personal’ violence rather than ‘impersonal’ violence at a distance (e.g., pushing a person in front of a runaway trolley, killing them and saving five people versus flicking a switch to divert the trolley to kill one person instead of five). Greene et al. argue that personal violence as a means to a utilitarian end, elicits a negative emotional response from the observer, and this
emotional response leads to an action being regarded as unacceptable. When *impersonal* violence is used to achieve the same utilitarian outcome, the negative emotional response to the violence is not as great, and so controlled processes dominate the decision, resulting in utilitarian considerations (the number of lives saved) having the greatest influence on judgement. This interpretation is supported by fMRI evidence that personal dilemmas increase neural activity in areas of the brain associated with emotion, such as the posterior cingulate cortex, the medial prefrontal cortex, and the amygdala (Greene et al., 2001; Greene 2007; but see McGuire, Langdon & Cotheart, 2009 and Greene, 2009).

Elaborating on the role of non-deliberative processing in moral cognition, heuristic theories state that moral judgements are often made using cognitive shortcuts, or rules of thumb (Sunstein, 2005). A heuristic is a decision making strategy that is generally quicker and uses less information than other (regression or Bayesian) strategies of decision making (Gigerenzer & Gaissmaier, 2011). In the case of moral judgement, Sunstein (2005) argues that people use heuristics such as, ‘do not knowingly cause a human death’ and ‘people should not be permitted to engage in moral wrongdoing for a fee’. These usually work well; however, when applied to new and perhaps unusual situations they can lead to nonsensical or poor decisions. For example, the heuristic ‘people should not be permitted to engage in moral wrongdoing for a fee’ ordinarily works well, as in the judgement that people should not be allowed to pay money to assault another person. However, the same heuristic applied to an emissions trading scheme might lead to a self-defeating conclusion (if the ultimate goal was to reduce pollution). Gigerenzer also argues that heuristics influence moral judgements and behaviours, but suggests that there is no need to distinguish moral from non-moral heuristics; the important heuristics that guide behaviour are the same in both moral and non-moral contexts. For example, ‘imitate your peers’ or ‘tit-for-tat’ – a description of a tendency people have to cooperate in the first encounter and to imitate their partner’s last behaviour in subsequent encounters – function well regardless of the decision domain (Gigerenzer, 2010).

In an approach focusing on how moral information is organized rather than how

---

1 Sunstein’s criteria for a bad decision or judgement seem to be based upon a consequentialist and utilitarian perspective. Though it may be that by objecting to emissions trading more pollution will occur, not monetizing wrongdoing can be an end in itself.
it is processed Gray, Young, and Waytz (2012) propose that all moral judgements are seen as an interaction between a moral dyad of an agent causing harm, and a moral patient suffering harm. This conceptual template influences perception and interpretation of potentially moral interactions, for example, witnessing someone suffering (a moral patient) leads to the completion of the dyad by inferring that a moral agent is likely responsible for the suffering, which may in turn influence moral judgements of the person identified as the moral agent.

*What do psychologists think about the moral good?*

Though theories of moral judgement are diverse, what they have in common is that they make claims about moral judgement using evidence almost exclusively from studies looking at moral transgressions and dilemmas. Empirical studies examining judgements of moral goodness are rare, and are largely limited to characteristics of morally good people (Aquino & Reed, 2002; Lapsley & Lasky, 2001; Walker & Pitts, 1998). For example, in two different studies, participants generated traits of someone who has good character (Lapsley & Lasky, 2001), or who is highly moral (Walker & Pitts, 1998). In follow up studies participants then identified which of these traits are most typical of a highly moral person, or a person with good character. Looking at the typical characteristics that appear in both studies\(^2\), the picture of a morally good person is someone who is honest, trustworthy, faithful, loyal, reliable, responsible, sincere and respectable. Aquino and Reed (2002) took a slightly different approach with a content analysis on the traits that participants identified as being associated with a moral person, which characterized this person as caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind. The moral individual identified by Aquino & Reed is a kinder and more compassionate person than the trustworthy and responsible individual that emerges from Lapsley and Lasky (2001) and Walker and Pitts (1998).

What constitutes a morally good action has received less attention. In the few studies on this question, participants typically read vignettes depicting an individual’s actions and the participant then evaluates whether the person has acted morally well (e.g., Gong & Medin, 2012; Takahashi et al., 2008). The specific actions participants

\[^2\] I took the most typical characteristics identified in each study and then list those characteristics that appear in both studies.
evaluate in these studies come from either the experimenter’s intuition or previous research. Given how few studies have examined morally good actions, and the eclectic nature of the examples of good actions used in these studies, it would be difficult to identify important aspects of the moral good from an examination of examples used in the psychological literature.

Though not directly concerned with judgements of morally good actions, the prosocial literature examines conditions under which people are likely to act prosocially, which could be thought of as an investigation concerned with, in part, when and why people perform morally good actions. A common definition of prosocial behaviour is of actions that are intended to benefit others, and not oneself (Batson & Powell, 2003; Twenge, Baumeister, DeWall, Ciarocco, & Bartels, 2007). Similarly, altruism has been defined as a motivation or desire to benefit someone else for his or her own sake rather than one’s own (Batson, 2011). As well as implying that a prosocial or altruistic action will likely benefit another person, these definitions also identify the intent of the person performing the action as central to the definition.

Perhaps as important as the outcomes of actions, for judgements of moral goodness, are the motives or reasons people have for their actions (Guglielmo et al., 2009). For example, the motivation to act in the interest of others is potentially as important for goodness judgements as the outcome benefitting others. Several studies have found that an agent’s motive can influence goodness judgements (Baird & Astington, 2004; Janoff-Bulman et al., 2009), for example, Janoff-Bulman et al. found that performing a good action because of an intrinsic desire to perform the action resulted in different goodness judgements than if the action was performed out of a sense of duty.

Beyond the few studies that have specifically looked at goodness judgements of actions, we might also look to the broader moral judgement literature to understand how people decide if an action is morally good. Moral judgement research typically examines the nature of moral judgements by examining how people evaluate moral transgressions and dilemmas. By examining transgressions and dilemmas and subsequently making claims about moral judgements in general, the assumption is that both judgements of moral goodness and wrongdoing are aspects of moral judgement, and that findings relevant to one type of moral judgement should generalize to all types
of moral judgement. This assumption can be seen in the footnotes or parenthetical remarks often found at the end of studies looking at moral judgement, suggesting that the findings likely apply to judgements of morally good actions as well (Gray et al., 2012).

Though the collective term morality covers both morally good and bad actions, and there is a perception that good and bad are opposites, and that actions range on a continuum from good to bad, should we expect findings relevant to judgements of morally bad actions to generalize to judgements of morally good actions?

Transgressions and dilemmas

Claims about moral judgements are potentially problematic if the empirical evidence used to support these claims comes exclusively from transgressions and dilemmas. Moral transgressions involve an agent performing a morally wrong or blameworthy action (Haidt, Koller, & Dias, 1993; Moll, de Oliveira-Souza, Bramati, & Grafman, 2002; Schnall, Haidt, Clore, & Jordan, 2008). Transgressions have been used to identify what kinds of actions are morally wrong, as well as how particular variables influence moral judgements. For example Schnall et al., (2008) provided participants with scenarios, and then asked them how moral or immoral the actions described in the scenarios were (1-extremely immoral, 7-perfectly ok).

Dilemmas are situations where competing moral principles are placed in opposition (Broeders, van den Bos, Müller, & Ham, 2011). For example in the classic trolley problem (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001) there is a runaway trolley that will kill five people. An observer has the option of flicking a switch to divert the trolley to a track where one person will be killed instead of five. The dilemma here involves placing in opposition principles of not harming another person (switching the tracks will kill someone who would not have been killed if the switch hadn’t been thrown) and the utilitarian principle, that the needs of the many outweigh the needs of the few. Dilemmas are used in studies examining variables that influence moral judgements (Cushman, Young, Hauser, 2006; Greene et al., 2001; Tassy et al., 2012; Valdesolo & Desteno, 2006) as well as in studies looking at moral reasoning and justification for moral judgements (Kohlberg, 1969; Zarinpoush, Cooper, & Moylan,
2000). In moral judgement studies using dilemmas, participants are usually asked if a particular approach to the dilemma is acceptable or unacceptable (Tassy et al., 2012), appropriate or inappropriate (Greene et al., 2001; Valdesolo & Desteno, 2006), forbidden, permissible, or obligatory (Cushman, Young, & Hauser, 2006). In responding to these questions participants are identifying whether a course of action is morally wrong or not. As with transgressions, these kinds of questions identify what participants think is morally bad, but not what is morally good.

Results from transgression and dilemma studies could confidently be generalized to judgements of goodness only if these two kinds of judgements are made using similar criteria or similar cognitive-affective mechanisms. There is evidence however that this is not the case.

Two kinds of judgements

There is strong evidence that negative events, emotions and information, influence behaviour and judgements to a greater degree than positive events (Baumeister, Bratslavsky, Finkenhauer, & Vohs, 2001). Specific to moral judgements, there is evidence that negative information has greater influence on impression formation (Baumeister et al., 2001), and is more revealing of a person’s character (Ybarra, 2002), which can subsequently influence moral judgements of actions (Nadler & McDonnell, 2011, Pizarro, Laney, Morris, & Loftus, 2006). There is also evidence that negative events increase motivation to search for meaning and to find out why an outcome occurred (Weiner, 1985) and that the resulting explanation will strongly influence moral judgements (Guglielmo, Monroe, & Malle, 2009). Furthermore negative mood and emotion elicited by morally blameworthy actions can lead to more systematic processing of information (Forgas, 1995), which then may result in explanations for negative actions having a greater influence on moral judgements than explanations of positive actions.

In a study explicitly looking at differences in how judgements of good and bad actions are made, Wiltermuth, Monin and Chow (2010) examined whether the tendency to praise others for positive moral behaviours correlates with the tendency to condemn others for negative behaviours. Participants were given vignettes that described both
morally good and bad behaviours and participants rated the behaviours from extremely immoral to extremely moral. The results indicate that ratings of the morally good and bad actions were uncorrelated, and factor analysis identified that moral judgements of good and bad behaviours loaded on different factors. Furthermore blame and praise judgements were predicted by different sub-scales of Aquino and Reed’s Moral Identity Scale (2002), with internalization (importance of moral traits for self) sub-scale predicting ratings of negative behaviours, and symbolization (doing things that express moral characteristics) predicting ratings of positive behaviours. These results suggest that moral judgement of good and bad actions use different criteria, and should not be treated as a continuum. Though not looking at moral judgments, related behavioural studies have generally failed to find a correlation between prosocial and antisocial behaviours (Krueger, Hicks & McGue, 2001).

Relatedly, Janoff-Bulman, Sheikh and Hepp, (2009) propose that separate motivational systems are associated predominantly with judgements of morally good and bad actions. The dual regulatory system of approach and avoidance motivations has a central role in psychology (Gable, Reis, Elliot, 2003; Higgins, 1998). Janoff-Bulman et al. (2009) propose that the avoidance or prescriptive system, which is based on behaviour inhibition, and is closely associated with negative outcomes, focuses on what we should not do, which mostly involves judgements of moral wrong doing. By contrast the approach or prescriptive system, which is based on behaviour activation and associated with positive outcomes, focuses on what we should do, and is associated mostly with performing morally good actions. The authors theorize that consistent with finding throughout psychology that people give more importance to negative events and outcomes (Baumeister et al., 2001), prescriptive judgements are likely to be harsher and more demanding than prescriptive judgements. Janoff-Bulman et al. (2009) found that participants rated prescriptive statements and vignettes (e.g., treat others fairly, donate to charity) as equally morally important as prescriptive statements and vignettes (e.g., do not cheat, do not steal), however prescriptive vignettes were judged to be more mandatory and less a matter of personal preference than prescriptive vignettes. It has been argued that an important factor that distinguishes moral from non-moral judgements is that moral actions are more likely to be judged as mandatory objective facts than conventional standards and personal preferences. Janoff-Bulman et al.’s findings
identify significant divergence between obligation to perform morally good and bad actions, which is a dimension that is one of the defining qualities of a moral judgement.

Takahashi et al., (2008) provide neuroimaging evidence for separate neuro-cognitive systems responsible for judgements of the moral good and bad. They found that different regions of the brain were associated with (in their words) judgements of morally beauty and depravity. FMRI scans indicated that judgments of morally good actions were uniquely associated with activity in the orbitofrontal cortex, while judgements of morally bad actions were uniquely associated with activity in the posterior superior temporal sulcus. This evidence suggests that different parts of the brain are recruited for judgments of morally good and bad behaviours (but see Young, Scholz, & Saxe, 2011). Interestingly other researchers have found evidence that judgments of different kinds of moral transgressions are associated with different neural systems, and are perhaps not a unified category. In an fMRI study Parkinson, Sinnott-Armstrong, Koralus, Mendelovici, McGeer, and Wheatley (2011) found evidence for dissociable neural systems associated with different kinds of transgressions. The authors identified that different areas of the brain showed more activity when people evaluated transgressions involving harm, dishonesty and disgusting transgressions, which they interpreted as evidence for separate cognitive systems associated with each type of transgression.

Conclusion

In sum, there is compelling evidence that judgements of morally good and bad actions are sufficiently different that research on the former cannot be automatically generalized to judgments of the latter. Given this assessment, and the scarcity of research looking at judgements of good actions, I propose taking an exploratory, inductive look at what a morally good action is, and what features of a morally good action are important for goodness judgements of actions. In the next chapter, I examine the folk concept of the good. Chapters 3 and 4 then empirically test the influence of several important factors, identified in the first study, on judgements of morally good actions, by presenting moral scenarios in which those factors vary systematically. In Chapter 5, I use the findings in Chapters 2 to 4 to reconsider previous empirical work on diminished capacity and the moral good.
Chapter 2

The folk concept of moral goodness

Research indicates that judgements of moral transgressions are different from judgements of morally good actions. A handful of studies have specifically looked at judgements of morally good actions but it is still unclear from the literature what a morally good action is. In this chapter I take an exploratory look at the folk concept of moral goodness.

In Chapter 1, I argued that one should be cautious generalising from research on judgements of moral transgressions and dilemmas, to judgements of moral goodness. Given that there has been little research examining perceptions of moral goodness and what constitutes a morally good action, I begin by considering what it means for something to be morally good.

Folk concepts

A difficulty with concepts such as the moral good is that they come from the ordinary world of language and people. Ordinary concepts resist strict definition and are difficult to pin down for examination (Gray, Young, Waytz, 2012; Medin, Wattenmaker, & Hampson, 1987; Wittgenstein, 1958). Understanding such ordinary or folk concepts\(^3\), however, can serve several purposes. Firstly we may want to understand exactly how people use a folk concept. For example Malle and Knobe (1997) wanted to understand under what conditions an action is identified as intentional.

Secondly understanding the folk concept can also help clear up confusion. Researchers can mistakenly identify only certain aspects of a concept if they just rely on their own understanding of it. Prior to the research of Malle and Knobe (1997), for

\(^3\) Also referred to as a prototype Lapsley and Lasky (2001) and as a vernacular concept, Griffiths, Machery, and Linquist (2009).
example, researchers had proposed numerous definitions of intentionality. An analysis of the folk concept allowed Malle and Knobe to reconcile these definitions, which, it turned out, focused on different aspects of the folk concept. Griffiths, Machery, and Linquist (2009) identified that researchers consistently misunderstood the folk concept of innateness, and Monroe and Malle (2010) were able to demonstrate that people do not have the conception of free will – one inconsistent with a determinate universe – that researchers believed them to have. A further benefit of a clearer understanding of the concept being investigated is that it allows researchers to develop more precise and specific hypotheses and to make more nuanced predictions.

Even when researchers are not specifically interested in folk concepts, they can still run into difficulty by ignoring them. Folk concepts are almost always broader and more nuanced than technical definitions (Griffiths, 1997). When a folk concept differs from a technical definition this can lead to inadvertent and inappropriate generalisation of findings, by applying the technical definition to situations that the technical definition was not meant to cover. This can cause confusion when researchers move back and forth between technical definitions and common usage (Griffiths, 1997; Resnik, 2000). Furthermore, as a practical matter, research participants may not be aware of the technical definition of a concept, and if researchers do not understand how participants use the folk concept then the answers participants give to the researcher’s questions will be misleading.

The folk concept of moral goodness

As in the examples above, examining the folk concept of moral goodness will help identify what it is that is being researched when we are looking at the moral good. It will help avoid unknown factors associated with the concept from influencing studies, and will allow for a better understanding of how participants answer questions about moral goodness. As well as a descriptive account of the folk concept of moral goodness, examining the folk concept will help reveal aspects of the concept that are directly relevant to how people make goodness judgements.

The picture of the moral good in the literature is of positive evaluation of actions and character with respect to moral standards. From the literature we can see that
judgements of morally good actions are influenced broadly by two things: the actions performed by an agent (Gong & Medin, 2012; Takahashi et al., 2008), and assessment of the mental states of an agent performing the actions (Baird & Astington, 2004; Janoff et al., 2009, Pizarro, Uhlmann, & Salovey, 2003). However, specifically which actions and mental states are relevant and most important is left unstated in this research. Study 1 takes an exploratory look at the folk concept of moral goodness to help fill in this gap in the research.

Study 1

This study will serve as an initial exploratory look at the folk concept of moral goodness, following a strategy employed by Malle and Knobe (1997, study 2). In that study, the researchers examined intentionality, which had been inconsistently defined in the literature, by asking participants “When you say that somebody performed an action intentionally, what does this mean?” The authors then used the factors identified as independent variables in subsequent studies to confirm that they did in fact influence judgements of intentionality.

Though I am primarily interested in goodness judgements of actions, the broader concept of moral goodness is examined to assess participants’ intuitions about moral evaluation of character as well. The importance of the relationship between moral evaluation of character and actions will be investigated in greater detail in later chapters.

Method

Participants

One hundred and eight (36 male, 72 female; \( M_{age} = 19.75 \)) first and second year psychology students at Otago University gave informed consent before participating in return for course credit.

Procedure and Materials

This study was run as the second of three unrelated studies completed by the same participants during a one-hour session. Participants were given a pen and a piece of
paper with the following instructions: “In this next task we are interested in your thoughts on morality. When you say that something is morally good, what do you mean? Please explain.” Nine lines were included after the question for participants to write their response (Appendix A). The amount of space participants were given to respond was selected to indicate that we wanted a relatively brief response. The question posed was left intentionally broad to allow for responses related to both actions and character.

Results and Discussion

Content analysis was conducted and categories were coded and analysed using the guidelines outlined in Krippendorff (2004). Specifically, I first read through all responses to formulate categories that accounted for the full range of responses (see Appendix B for coding instructions). Twenty categories, within five superordinate themes, were identified (see Table 2.1 and Discussion).

---

4 Participants were asked “What do you mean” to encourage participants to reflect on what they themselves say and to encourage participants not to recite formal moral theories they may be aware of and not to list what they think other people think it means.
<table>
<thead>
<tr>
<th>Theme</th>
<th>Category</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonyms</td>
<td>Described the moral good using a synonym for the good such as good, right, just and ethical.</td>
<td>“it's really hard to describe, when something is morally good it's when it's doing the right thing”</td>
<td></td>
</tr>
<tr>
<td>Target of morality</td>
<td>Character</td>
<td>Described a person as the target of moral evaluation.</td>
<td>“someone who is morally good is”</td>
</tr>
<tr>
<td>Action</td>
<td>Described an action as the target of moral evaluation.</td>
<td>“if you do something morally good you have acted ...”</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>Benefit others</td>
<td>An action or person should benefit others.</td>
<td>“the thing that ... benefits the greatest number of people”</td>
</tr>
<tr>
<td></td>
<td>An action or person should not harm or negatively affect others.</td>
<td>“is something that doesn’t negatively affect others”</td>
<td></td>
</tr>
<tr>
<td>No benefit to self</td>
<td>For an action or person to be morally good a person should not desire to benefit from their actions.</td>
<td>“doing something good that doesn’t involve personal gain”</td>
<td></td>
</tr>
<tr>
<td>Motives</td>
<td>Principle</td>
<td>An action or person is morally good if someone acts in accordance with moral codes, rules, standards, or principles.</td>
<td>“it doesn’t conflict or contradict the ethical principles I attempt to live my life by”</td>
</tr>
<tr>
<td>Other interest</td>
<td>Should be motivated by a desire to benefit or act in the interest of others</td>
<td>“it is good when it is done with good intentions”</td>
<td></td>
</tr>
<tr>
<td>No motive to harm</td>
<td>A person should not have a motive to harm or negatively affect others.</td>
<td>“an act which may benefit others and not purposefully cause harm”</td>
<td></td>
</tr>
<tr>
<td>No self interest</td>
<td>A person should not desire to benefit from their actions.</td>
<td>“do something not for the good of yourself”</td>
<td></td>
</tr>
<tr>
<td>Sympathy</td>
<td>A person feels or a good action is motivated by empathy, sympathy or compassion for others.</td>
<td>“feels compassion”</td>
<td></td>
</tr>
<tr>
<td>Justification</td>
<td>Society</td>
<td>Participants mentioned that morality is decided by societal rules of conduct.</td>
<td>“it fits the societal beliefs of what is right”</td>
</tr>
<tr>
<td>Self</td>
<td>Morality is justified by a person’s own sense of what is right, or personal beliefs.</td>
<td>“an individual’s morals differ from person to person”</td>
<td></td>
</tr>
<tr>
<td>Religion</td>
<td>Indicated that something is morally good because a religion prescribes the rules.</td>
<td>“something that is not frowned upon in the bible”</td>
<td></td>
</tr>
<tr>
<td>Subjective</td>
<td>Participant explicitly mentioned that morality is subjective.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective</td>
<td>Participant explicitly mentioned that morality is objective.</td>
<td>“adheres to some (wholly or partially) objective, external stand of morals”</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Feel good not regret</td>
<td>That acting in a morally good way will make you feel good and you will not regret your actions later.</td>
<td>“is not ... anything that should leave you feeling guilty or ashamed”</td>
</tr>
<tr>
<td>Feeling</td>
<td>Moral judgement are evaluated by a feeling that an action is good or bad.</td>
<td>“using your own inner ‘sense’ of what is the correct way to behave”</td>
<td></td>
</tr>
<tr>
<td>Strong belief</td>
<td>Moral beliefs are something a person strongly believes in.</td>
<td>“something that you strongly believe in”</td>
<td></td>
</tr>
<tr>
<td>No obligation</td>
<td>Indicated that a person is not obligated to perform morally good actions.</td>
<td>“an act which is done … not because you are obliged or expected to”</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Category codes, with brief description and illustrative example taken from participants’ responses. The categories are organized into five major themes that emerge from participants’ responses.

**Coding**

To evaluate the validity of the categories, two coders were trained to categorise participants’ responses into the categories. The coders’ task was to read each of the participants’ responses and identify which categories the participants’ response fell into. Coders were given written and oral instructions (Appendix B), and coded several practice responses and discussed any ambiguities with me before they began on the full
There was no limit to the number of categories to which a response could be assigned, and a single sentence could contain several ideas that would then be coded into several categories. Coders coded the data independently, with differences resolved later by discussion. The coders identified from the responses of 108 participants 435 categorizable ideas. Participants’ responses were coded into between 2 and 7 categories with a mean of 4.03 categories per participant. Table 2.2 shows the percentage of participants who mentioned a category, the percentage agreement coders had for responses in each category, as well as Krippendorff’s alpha.

For most categories Krippendorff’s alpha was below or just above the minimum reliability required to draw any sort of inference. Suggested $\alpha$ values for tentative conclusions are of .667 and .8 for acceptable reliability (Krippendorff, 2004), and the 95% confidence intervals for $\alpha$ span much of the possible range. Percent agreement levels were quite high, for a comparison Malle and Knobe (1997) reported 88% agreement initially which rose to 95% for their categories relevant to intentionality. In this study the lowest percent agreement item was 84.76% with 9 of the 25 exceeding 95%. However, percent agreement has been shown to be a poor measure of inter-coder reliability (Krippendorff, 2004).

Coders reported, anecdotally, that disagreement was most often due to overlooking a relevant category – not surprising given the difficulty of simultaneously using 20 categories to code any given aspect of a participant’s response. In 13 cases (3%) the coders could not resolve the conflicts, in most cases because they could not agree whether “believing” something represented a belief or a principle. Coders also mentioned some difficulties in determining whether participants were indicating that morality was justified by the self or society, and distinguishing motives from outcomes.

Given the low coder reliability of the categories any inferences made from the descriptive statistics alone are highly speculative. However this study was designed to be an initial look at the folk concept of moral goodness and to help inform hypothesis generation in subsequent experiments, and so for this purpose of hypothesis generation low reliability is not too much of a concern.
<table>
<thead>
<tr>
<th>Theme (Present %)</th>
<th>Category</th>
<th>Present %</th>
<th>Agreement %</th>
<th>$\alpha$</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonyms</td>
<td></td>
<td>31.40</td>
<td>85.71</td>
<td>.701</td>
<td>.541</td>
<td>.855</td>
</tr>
<tr>
<td>Target of morality (100)</td>
<td>Action</td>
<td>94.40</td>
<td>93.33</td>
<td>.596</td>
<td>.222</td>
<td>.859</td>
</tr>
<tr>
<td></td>
<td>Character</td>
<td>17.60</td>
<td>92.38</td>
<td>.789</td>
<td>.620</td>
<td>.937</td>
</tr>
<tr>
<td>Outcome (50.00)</td>
<td>Benefit others</td>
<td>36.10</td>
<td>84.76</td>
<td>.684</td>
<td>.523</td>
<td>.813</td>
</tr>
<tr>
<td></td>
<td>No harm</td>
<td>22.20</td>
<td>94.29</td>
<td>.839</td>
<td>.690</td>
<td>.944</td>
</tr>
<tr>
<td></td>
<td>No benefit to self</td>
<td>4.63</td>
<td>89.52</td>
<td>.421</td>
<td>.060</td>
<td>.726</td>
</tr>
<tr>
<td>Motives (68.50)</td>
<td>Principle</td>
<td>31.40</td>
<td>84.76</td>
<td>.663</td>
<td>.514</td>
<td>.823</td>
</tr>
<tr>
<td></td>
<td>Other interest</td>
<td>28.70</td>
<td>88.57</td>
<td>.721</td>
<td>.561</td>
<td>.854</td>
</tr>
<tr>
<td></td>
<td>No self interest</td>
<td>18.50</td>
<td>91.43</td>
<td>.613</td>
<td>.350</td>
<td>.840</td>
</tr>
<tr>
<td></td>
<td>No motive to harm</td>
<td>7.40</td>
<td>92.38</td>
<td>.373</td>
<td>-.207</td>
<td>.7806</td>
</tr>
<tr>
<td></td>
<td>Sympathy</td>
<td>4.60</td>
<td>99.05</td>
<td>.885</td>
<td>.636</td>
<td>1.000</td>
</tr>
<tr>
<td>Justification (53.70)</td>
<td>Society</td>
<td>37.00</td>
<td>86.67</td>
<td>.735</td>
<td>.593</td>
<td>.871</td>
</tr>
<tr>
<td></td>
<td>Self</td>
<td>20.40</td>
<td>95.24</td>
<td>.789</td>
<td>.620</td>
<td>.937</td>
</tr>
<tr>
<td></td>
<td>Religion</td>
<td>10.20</td>
<td>97.14</td>
<td>.827</td>
<td>.577</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Subjective</td>
<td>13.00</td>
<td>97.14</td>
<td>.906</td>
<td>.754</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Objective</td>
<td>1.90</td>
<td>98.10</td>
<td>.001$^5$</td>
<td>-1.000</td>
<td>.529</td>
</tr>
<tr>
<td>Other (16.67)</td>
<td>Feel good no regret</td>
<td>9.30</td>
<td>99.05</td>
<td>.942</td>
<td>.819</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Feeling</td>
<td>9.30</td>
<td>97.14</td>
<td>.809</td>
<td>.599</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Strong belief</td>
<td>2.80</td>
<td>98.10</td>
<td>.796</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>No obligation</td>
<td>1.90</td>
<td>98.10</td>
<td>.659</td>
<td>.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note. Present%: The percentage of participants with responses of each type, Agreement% is the percent agreement between coders for each category, e.g. 85.71% of instances where coder 1 identified a synonym coder 2 also identified a synonym, Krippendorff’s alpha ($\alpha$), 95% confidence interval for $\alpha$. Categories are grouped to match the thematic grouping of the discussion.

$^5$ Krippendorff’s alpha apparently suffers the same problems as Cohen’s kappa for situations of very high agreement (Cicchetti & Feinstein, 1990), i.e. it falls apart.
Synonyms

Participants’ responses broadly indicated concerns with outcomes, motives, justifications and the nature of goodness and features of moral judgements. About one third (31.40%) of participants explained what they meant by moral goodness with references to synonyms, such as, “good”, “right”, “just” and “ethical”. Though distinctions can be made between these terms, it was decided that this preliminary analysis was not sensitive enough to be able to discriminate meaningfully between these ideas.

What can be morally good? Targets of morality

All participants included a target of moral judgements in their response. Nearly all participants (94.40%) indicated that an action could be the target of moral evaluation (e.g., “if you do something morally good you have acted...”); “when something is morally good, this refers to a person’s actions”). Indeed, most research on moral judgement considers actions, and the moral goodness of actions is also a main focus of this thesis.

In addition, a substantial minority of participants (17.60%) indicated that the target of moral evaluation was a person (e.g., “when someone has morals they are able to differentiate between right and wrong”), a response I termed Character. Moral judgements of actions are the main focus of the moral judgement literature however character judgements are potentially more frequent given the automatic nature of spontaneous trait inferences (Overwelle et al., 2012).

There is evidence in the literature that moral judgements of actions are not just judgements of actions but also of the person performing them. Pizarro and Tannenbaum (2012) argue that given how people automatically evaluate the moral character of others (Fiske, Cuddy, & Glick, 2007; Willis & Todorov, 2006) these evaluations are likely to influence subsequent moral evaluations of actions, and that people evaluate actions by asking, “Is this the kind of action a good or bad person would perform?” Furthermore Pizarro and Tannenbaum (2012) argue that findings in the literature on moral transgressions and blame that are interpreted as processing biases (Alicke, 2000) can more parsimoniously be explained as the persistent influence of character judgements on action judgements.
Outcomes
Fifty percent of participants identified that good actions were associated with particular outcomes, which is consistent with research suggesting the importance of outcomes for goodness judgements (Gong & Medin 2012; Janoff-Bulman et al., 2009; Takahashi et al., 2008). Three specific outcomes were mentioned that were relevant to something being morally good: the action benefits other people or furthers their interests (Benefit others); the action doesn’t harm others (No harm); and the action does not benefit the person performing the action (No benefit to self).

Responses categorised as Benefit others, mentioned by 36.1% of participants, indicated someway in which the outcome benefits others or furthered their interests (e.g., “has a positive impact on people”; “the thing that ... benefits the greatest number of people”). This outcome description is consistent with behaviour-centric definitions of prosocial behaviour as voluntary behaviour that benefits others (Eisenberg & Miller, 1987; Malti & Krettenaur, 2012) as well as behavioural accounts of altruism. For example Fehr (2003) defines altruism as costly acts that confer economic benefits on other individuals.

Responses categorised as No benefit to self, mentioned by 4.6% of participants, indicated that a person should not benefit from their actions (e.g., “doing something good that doesn’t involve personal gain”). Self interest is central to ideas of morality and it has been argued that the function of moral systems is to regulate selfishness to make social life possible (Haidt, 2008).

Responses categorised as No harm, mentioned by 22.2% of participants, indicated that a good action should not harm others, (e.g., “is something that doesn’t negatively affect others”; “... without negatively impacting anyone or anything”). Harming others has been demonstrated in numerous studies to be a morally blameworthy action (Haidt & Graham, 2007) and is a staple outcome of most moral transgression research.

Motives
Nearly seventy percent of participants identified at least one motive associated with moral goodness. Motives have been identified as central to moral judgements but have
received relatively little discussion in the broader morality literature (Baird & Astington, 2004; Guglielmo et al., 2009).

Responses categorised as Principle, the most frequently mentioned motive for performing a good action (31.40% of participants), indicated that something is good if someone acts in accordance with moral codes, rules, standards, or principles, (e.g., “it doesn’t conflict or contradict the ethical principles I attempt to live my life by”; “it adheres to some ... standard of morals”). Acting according to a principle as opposed to a desired outcome is captured by the philosophical distinction between rule following (deontological ethics) and concern for outcomes (consequentialism) (Greene, 2007). Batson, Ahmad and Tsang (2002; see also Batson & Powell, 2003) identified a motivation to act in accordance with a principle as one of the four motivations for prosocial behaviour.

Responses categorised as Other interest, mentioned by 28.70% of participants, indicated the importance of a desire to benefit or act in the interest of others (e.g., “it is good when it is done with good intentions”; “something that you think helps others”). Though prosocial actions have been defined as actions that benefit others, more commonly the intent of the action is taken into account in the definition. For example, prosocial behaviour has been defined as “actions intended to benefit one or more people other than oneself” (Batson & Powell, 2003, p.463) and “prosocial behaviour is performed to benefit others, rather than to benefit the self (Twenge et al., 2007, p.56). Altruism captures a similar idea to prosocial behaviour but is mostly used to describe a particular desire (or motivation) to benefit someone else for his or her sake rather than that of the actor (Batson, 2011).

Responses categorised as No self interest, mentioned by 18.50% of participants, indicated that for a person or action to be morally good, a person should not desire to benefit from their actions, (e.g., “something done unselfishly”; “do something not for the good of yourself”). This motive appears as the second part of many of the descriptions of prosocial and altruistic behaviour indicating that for an action to be considered prosocial people should not only have a motive to benefit others but they should also not have a motive to benefit themselves. Lin-Healy and Small (2012) found that when people infer selfish motivations for donations to charity, the donor is evaluated as less nice, altruistic, kind and generous.
Responses categorised as *No motive to harm*, mentioned by 7.40% of participants indicated that a person should not have a motive to harm, or negatively affect others, (e.g., an act which may benefit others and not purposefully cause harm). As noted above, harm is inextricably linked to moral transgressions and it follows that a motive to harm would be incompatible with a good action as well. Intent to harm has been shown to increase how morally wrong an action is judged to be (Cushman, 2008). It is noteworthy that participants mentioned avoiding harmful outcomes more frequently than motivations to avoid harmful outcomes, whereas the reverse was true for the presence of beneficial outcomes. That is, participants appear to indicate that a moral action involves the motivation to benefit others, regardless of whether that benefit eventuates, and the absence of harm, regardless of whether that harm is intended. Very few participants mentioned that people should not have a motive to harm others, possibly because this is not a motive that a person ever has for performing a good action. Harming others can be a side effect, or a means to achieve some potentially good end; however, being motivated to harm others and harming others as a consequence of trying to do some other good are two different situations.

Finally a small number of participants’ responses (4.60%), were coded as *Sympathy*, indicating that a good person feels – or a good action is motivated by – empathy, sympathy or compassion for others. Empathy, sympathy and compassion are related concepts in common usage (Goetz, Kelter & Simon-Thomas, 2010; Hoffman, 2000), and have all, at different times, been described as resulting from the suffering of others (Batson, 1991, Goetz, Kelter & Simon-Thomas, 2010). Feeling sympathy or compassion for another person has been shown to increase prosocial motivation (Batson et al., 1997), and studies examining accounts of good character have all identified empathy or compassion as a characteristic of a good person (Aquino & Reed, 2002; Lapsley & Lasky, 2001; Walker & Pitts, 1998).

*Justifications*

In addition to identifying what features of an action or person make them good, fifty four percent of participants identified where moral rules and norms come from. Though people have different systems of justification (Goodwin & Darley, 2008) and vary in how they justify morality, moral judgements tend to be very similar across gender,
education, politics, and religion (Banerjee, Huebner & Hauser, 2010). Of the justification categories participants mentioned society the most frequently (37.0%), followed by the self (20.40%), followed by religion (10.20%)\(^6\). Responses were categorised into justifications from Society if participants mentioned that morality is decided by societal rules of conduct (e.g., “it fits the societal beliefs of what is right”; “that abides by social norms”). The category Self identified that morality is justified by a person’s own sense of what is right, or personal beliefs (e.g. “an individuals morals differ from person to person”; “it is related to what my personal morals are”), and Religion indicated that something is morally good because God, religion, or a sacred text prescribe the rules (e.g., “something that is not frowned upon in the bible”). Furthermore responses were categorised as Subjective (13%), or Objective (1.90%) if participants explicitly mentioned that morality is subjective or objective, respectively (e.g., “adheres to some (wholly or partially) objective, external standard of morals”).

Other responses

Although the majority of participants’ responses mentioned either the reasons for behaviour or its outcome, a smaller number of additional responses did not fit these categories (16.67%). A number of participants (9.3%) alluded to the emotional consequences of performing morally good and bad actions (labelled Feel good don’t regret). Participants indicate that performing a morally good action will result in the person performing the action feeling good and/or not regretting their actions (e.g., “is not ... anything that should leave you feeling guilty or ashamed”; “you would feel good about what you had done... and would not regret it”). Previous research has also identified the particular emotional and physiological consequences of doing something right and wrong. For example guilt and shame often follow morally transgressing (Baumeister, Stillwell, & Heatherton, 1994; Smith, Webster, Parrott, & Eyre, 2002).

The same proportion of participants indicated that judgements of goodness are made using a feeling that the judge has (e.g., “using your own inner ‘sense’ of what is the correct way to behave”). This is consistent with models of the role of emotion in moral judgement (Haidt, 2001; Schnall et al., 2008), which argue that moral judgements are made using emotion as a heuristic.

\(^6\) Based on other studies run on the same population, about 50% identify themselves with an organized religion.
Responses categorised as *Strong belief*, mentioned by 2.80% of participants, indicated that morality involves strong feeling or beliefs (e.g., “when you are in a situation that you use the saying ‘morally good’ you generally have strong feelings towards that situation.”; “something that you strongly believe in”). This is consistent with the literature identifying judgements of moral transgressions as more serious than conventional transgressions (Nucci, 2001; Sousa, 2009) and work indicating that moral beliefs and attitudes are generally stronger than non-moral beliefs and attitudes (Skitka, 2010).

Finally several participants (1.9%) indicated that a person is not obligated to perform morally good actions, categorised as *No obligation* (e.g., “an act which is done … not because you are obliged or expected to”). Though the transgression literature talks about conforming to moral rules as obligatory, Janoff-Bulman et al. (2009) presented evidence that performing morally good actions is generally considered to be less obligatory than not morally transgressing. This may be another important distinction between judgements of morally good and bad actions.

In sum this study provides a glimpse at the concept of moral goodness as it is ordinarily used. Participants identified that both people and their actions can be targets of moral evaluation and that particular actions and outcomes as well as particular reasons for acting are typical of a morally good action. Participants also identified justifications for why something or someone is morally good, as well as identifying a number of aspects of morality and features of goodness judgements.

**Limitations**

Low coder reliability in some cases limits conclusions that can be drawn from the data. Nevertheless, there were few items that coders identified as not fitting the existing categories. This suggests that the categories adequately encompass the range of responses from participants. It is likely however that the categories are too broad to identify all of the interesting variations within categories. As this study was only intended as an initial qualitative examination of moral goodness in a particular population, and as a means of hypothesis generation for subsequent studies, this limitation is not fatal. However, additional, and more focussed descriptive research is needed to fully characterise the full range of individuals’ responses to my question.
Relatedly, this picture of the moral good is limited by the homogeneous demographics of the sample. Participants were all university students of a similar age and background attending a New Zealand university. However given that there is evidence that moral judgements are similar across gender, education, politics, and religion (Banerjee, Huebner & Hauser, 2010), the major aspects of the concept identified by participants are likely to be similar across groups. In any case, given that most moral judgement research is conducted on a similar demographic, we still want to understand how this group, who are the participants in most psychological studies, thinks about moral goodness.

A further limitation was that though we identified that participants indicated either actions or character could be the target of moral evaluation this distinction was not maintained across other categories. For example a participant may have indicated that a good action is one where you help other people or a good person is someone who helps other people, in both cases the statements would have been coded as Benefit others. This could be misleading if we infer that one of the categories speaks to the concept of a good action when in fact the instances of that category all came from descriptions of character.

The folk concept and moral judgement

The aims of this study were to take an initial exploratory look at the folk concept of moral goodness and to identify aspects of the folk concept of a morally good action that are relevant to goodness judgements of actions, and to generate testable hypotheses for the empirical part of the thesis.

A number of promising factors were identified. Some, such as the importance of particular outcomes, are intuitive: that moral actions benefit others, do not cause others harm and do not benefit the person performing the action. The presence or absence of these outcomes are predicted to influence goodness judgements. Others are less obvious, and untested. For example, participants identified that a characteristic of doing something good is that you feel good after performing it. Perhaps the observation that a person felt good or morally uplifted by their actions would increase the perceived morality of their actions. Similarly, identifying that a person performed an action not because they were obligated to perform it, but because they wanted to, could also
increase goodness judgements.

However, perhaps the most important finding from this study was the importance participants placed on performing a morally good action for the right kind of reason. Motives have been found to influence moral evaluations of actions (Baird & Astington, 2004; Guglielmo et al., 2009), but beyond identifying that good and bad motives matter, very little has been said about how specific motives influence goodness judgements of actions. Given that 70% of participants identified motives as central to what it means for something to be morally good – which was more than for actions – further investigation into the role of motives in goodness judgements of actions seems warranted. In Study 2, therefore, I consider the causal influence that motives identified in this study have on goodness judgements.
Chapter 3
Motives and goodness judgements

Study 1 identified several motives that are central to the concept of a morally good action. This chapter follows up on the findings of Study 1 and experimentally examines the influence these motives have on observers’ goodness judgements of actions. The mediating role of character judgements between motive and goodness judgements of actions is also examined.

Study 1 examined the folk concept of moral goodness. One of the most important findings from that study is that the state of mind of the actor is a critical part of what makes an action morally good. In particular, participants identified several motives that they believed are the right (or wrong) reasons for performing a morally good action. In this chapter I examine experimentally the influence these motives have on an observer’s goodness judgements of actions. I begin by discussing how people explain the behaviour of others, the role motives play in explanations of behaviour, and the evidence for the influence of motives on moral judgement.

Mind is central to morality
Moral judgements are mostly reserved for things that are attributed minds (Gray, Gray, & Wegner, 2007; Guglielmo, Monroe, & Malle 2009; Knobe & Prinz, 2008), although people can morally evaluate non-minded things (Jarudi, Kreps, & Bloom, 2008). Research suggests that agency (the ability to make choices in light of mental states) is a characteristic of entities that can be morally evaluated (Gray et al., 2007). The ability of an entity to feel or experience is a characteristic of an entity that can be transgressed against (and perhaps a good deed done for; Knobe & Prinz, 2008; Gray et al., 2007).

Humans treat things with minds differently from other things in the world. Minds are understood using a framework of concepts (such as agent, intention, belief, desire), which is often referred to as theory of mind (Brune, 2005; Malle, 2004; Premack & Woodruff, 1978). This theoretic framework is used to predict the actions of minded things and to explain behaviours in terms of mental states, as well as to make moral evaluations (Knobe & Prinz, 2008).
The morality of an agent’s actions is closely linked with how the agent’s actions are explained (Guglielmo et al., 2009). Entities with minds differ from other sorts of things, in that their movements in the world can be explained by means of the ordinary mechanistic causal explanations used for non-minded things, and in terms of mental states, using theory of mind explanations of behaviour or folk psychology (Guglielmo et al., 2009; Malle, 2004).

Mind and behaviour explanations

Behaviour explanations answer questions about why a person did what they did. For example, “Why didn’t she reply to my email? Because she was busy” (Malle, 2004). Once it has been identified that there is an agent and an action, there is an immediate search to identify whether the action is intentional or unintentional (Malle, 2004).

Unintentional actions are actions that people do not mean to do, and are explained by causes, for example, knocking a cup of water onto the floor, being startled when one is surprised, or yawning (Malle, 2004). Intentional actions, on the other hand, are actions that people mean or choose to do in light of their beliefs and desires (Guglielmo et al., 2009; Malle, 2004). An intentional action involves an agent having an intention to act, and this intention is formed in light of the agent’s beliefs and desires, and while performing the action, the agent is aware of what they are doing and has the skill to perform it (Malle & Knobe, 1997). Unlike unintentional actions, intentional actions are explained using three kinds of explanations: reason explanations, causal history explanations, and valuings (Malle, 2004). Reason explanations use mental states such as beliefs and desires to explain behaviours; for example, a person may have bought some new clothes because they wanted (desired) to impress someone, or may have driven their car on a particular road because they thought (believed) that it was a quicker way to get home. The second type of explanation for intentional actions is a causal history explanation. Causal history explanations make reference to causal factors internal or external to a person that give rise to a belief or a desire. For example saying someone drank a glass of water because they were thirsty explains a behaviour by referring to a causal factor, thirst. Thirst is neither belief nor desire so it is not a reason

7 Though these factors are relevant to an action being intentional, there are degrees of intentionality, and each condition is neither necessary nor sufficient (Malle & Knobe, 1997).
explanation, but it provides an explanation for why someone might desire a drink of water. Finally, valuings are explanations that do not appeal to reasons or causes of reasons, but rather to evaluations that motivate a person’s behaviour; for example, I invited some friends over to eat this Saturday night, because I love hosting dinner parties (Malle, 2004).

**Behaviour explanations and moral judgement**

**Intention and control**

Research has found that intentional actions are judged more blameworthy or morally wrong than unintentional actions (A stington & Baird, 2004; Lagnado & Channon, 2008). Children use intentionality information, but their judgements are mostly influenced by the valence of the outcome (Zelazo, Helwig, & Lau, 1996). Intentionality is more important as children get older (Zelazo et al., 1996), and for adults intentionality is more important than outcomes for judgements of wrongdoing (Cushman, 2008; Young, Cushman, Hauser, & Saxe, 2007). There is also evidence that the importance of intent differs depending on the kind of moral transgression; for example, Young and Saxe (2011) found that intent was more important for judgements of harm violations than purity violations. Furthermore it has been shown that certain groups make less use of intent information; Callosotomy patients make less use of intentionality information than controls when making moral judgements (Miller et al, 2010), as do people with high functioning autism (Moran et al., 2011).

The degree to which a person is perceived to be in control of their actions and choices also influences judgements of their moral transgressions (Alicke, 2000; but perhaps not judgements of moral goodness, Pizarro, Uhlman, & Salovey, 2003). A person must have the capacity to form an intention in light of their beliefs to be held fully morally responsible for their actions. For example, someone suffering from brain damage or experiencing an extreme emotional episode may be judged to lack sufficient control over their actions to be completely blameworthy (Alicke, 2008; Guglielmo et al., 2009; Pizarro et al., 2003). The perceived inability to enact choice in the usual way is referred to as a loss of personal control (Alicke, 2000, 2008) or diminished capacity (Guglielmo et al., 2009). Whether diminished capacity influences judgements of blame
due to the effect it has on perception of intentionality or whether it influences moral judgements independently of intention is not known (Guglielmo et al., 2009).

Motives

If you witness someone hitting another person on the back, your evaluation of their actions will differ drastically depending on whether you think they hit them out of anger, or were trying to stop them from choking.

There is evidence that motives influence moral evaluation of intentional actions (Baird & Astington, 2004; Guglielmo et al., 2009; Nadler & McDonnell, 2011). Motive is used as a collective term for the mental states (beliefs, desires and valuings) that are used to explain intentional behaviour.  

Children as young as four can make use of information about motives when evaluating an agent’s actions, and the use of motives plays a larger role in moral evaluations as children get older (Baird & Astington, 2004). Almost all studies that examined the influence of motives on moral judgement have looked at whether the person being evaluated has a good, bad or neutral motive. Yuill (1984) examined how children make use of information about bad and neutral motives, and Nelson (1980) looked at how children make use of good and bad motives. In each study the researchers found that young children’s moral evaluations of another child were influenced by the child’s motives. However in both studies children were asked whether the boy in the story was a “good boy or a bad boy”, which is ambiguous as to whether the evaluations the children make are of the boy’s actions, the boy himself, or both.

Baird and Astington (2004) specifically looked at children’s moral evaluations of actions and examined good and bad motives for performing an action. Children were presented with stories in which a child had a good motive (e.g., they wanted to throw a

---

8 For example, Phoebe’s motive for borrowing the book is that she wants to read it on the train on the way home (desire). Phoebe’s motive for borrowing the book is that she thinks she will be able to read it on the train on the way home (belief). Phoebe’s motive for borrowing the book is that she loves reading on the train (valuing).

9 Motive is a general term, which does not fit precisely with the distinctions of explanations of intentional behaviour made by Malle (2004). The contents of causal history explanations that the agent is most likely not aware of when they are performing an action do not fit with what we normally think of as motives. For example, saying that her motive (or motivation) for calling her friend an idiot is that she is an angry sort of person, does not make sense; however, saying the reason she called her friend an idiot is that she is an angry sort of person fits with common usage. I discuss this further in chapter 4.
ball to a friend who didn’t have a toy to play with), or a bad motive (they threw a ball at the friend to try and hit them because they were angry with them). The authors found that children’s moral judgements differed depending on whether the child had a good or bad motive, though good and bad motives were compared with each other, so it is not clear whether good motives, bad motives or both influenced moral judgements.

Nadler and McDonnell (2011) examined motive in a legal context, and looked at how motives influence judgements of responsibility and blame for unintentional harm. They found that a bad motive resulted in more responsibility and blame attributed to the agent whose actions led to an accident, than if the agent had a good motive for their actions.

Though not specifically looking at moral judgements of actions, Reeder and colleagues have conducted a series of studies identifying the importance of motives for trait attribution (Reeder, 2009; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002; Reeder, Monroe, & Pryor, 2008; Reeder, Vonk, Ronk, Ham, & Lawrence, 2004). Reeder’s multiple inference model builds on classic person/situation models of trait attribution (Kelly, 1973), and argues that situational forces influence trait inferences through inference about a person’s motives (2009). Motives have been shown to be better predictors of traits than situation attributions (Reeder et al., 2004). Consistent with this model it has been found that spontaneous goal (or motive) inferences are made faster than spontaneous trait inferences (Overwalle, Duynslaeger, Coomans, & Timmermans, 2012). Furthermore Reeder et al. (2002) conducted a study that specifically showed motives influencing moral trait or character judgements. Though the trait literature does not look at moral judgements of actions per se, these studies show the importance of motives for evaluations of character, which research suggests has consequences for judgements of actions. I discuss the relationship between character and action judgements later in this chapter.

To the best of my knowledge the only study to compare the influence of specific motives on an observer’s goodness judgements of an agent’s actions is reported by Janoff-Bulman et al. (2009). The authors examined the influence of two separate motives: a simple desire to perform a good action and a motive to act in accordance with duty. The authors found that when the agent performed a morally good action by avoiding a transgression, the actions of the agent who acted out of duty (“Brian just got
his paycheque and would like to gamble, but decides he shouldn’t go to the track”]) were rated more morally good than the actions of the agent who acted out of a desire not to transgress (“Brian just got his paycheque and knows he could gamble, but decides he doesn’t want to go to the track”). When the agents actively performed a good action (for example in the desire condition: “Mary knows she could just walk past the homeless man, but wants to give him the money and hands him $2.00”), there was no difference between the motive conditions. However a simple comparison between the motives of desire and duty is difficult because the duty condition involved two separate desires, the agent had a desire not to perform the good action and at the same time felt they had a duty to do it, whereas the agent in the desire condition only had desire to perform the good action.

So there is evidence that motives are important for evaluations of morally good actions. However beyond identifying that good and bad motives influence goodness judgements differently, there has been little said about the role of motives for making goodness judgements (or moral judgements in general). In Study 1 participants mentioned more than any other factor that motives are characteristic of a morally good action, and yet the literature is mostly silent as to what specific motives are the right kinds – or wrong kinds – of motives for performing morally good actions. A few studies have examined the influence of motives on goodness judgements of actions, but these studies have either compared good to bad motives, which makes it difficult to identify whether the good or bad (or both) motive is influencing goodness judgements (Baird & Astington, 2004), or the study does not allow for a clear interpretation of the role of a specific motive on goodness judgements of actions (Janoff-Bulman et al., 2009). Identifying what specific motives are associated with morally good actions will help us to understand what it means for an action to be morally good, as well as what motives are likely to influence goodness judgements of actions.

There have not been any studies that have systematically examined what kinds of motives are important for performing a morally good action, and how they influence judgements of morally good actions. The difficulty perhaps in investigating motives is how to approach and organise motives in a useful way. A person can have an infinite number of motives for their actions and so a principled way of looking at motives is needed. I suggest that a good place to start is to look at those motives closely associated
with the concept of a morally good action. In Study 1 participants identified the right kinds of motives for performing morally good actions, as well as the wrong kinds of motives. As motives central to the concept of a morally good action these motives are likely to be the most influential and important motives for influencing goodness judgements of actions.

In Study 2, therefore, I propose to experimentally examine the motives identified by participants in Study 1 to see if they do indeed influence goodness judgements of actions. People do not always have insight into the reasons behind their judgements, actions (Gazzaniga, 2000; Nisbett & Wilson, 1977) and moral evaluations (Haidt, 2001, Haidt, Bjorklund, & Murphy, 2000; Wheatley & Haidt, 2005). Nisbett and Wilson (1977) argue that people do not have true insight into their cognitive processes, but rather rely on theories of how stimuli influence their judgements and actions. The authors conducted a number of studies in support of this argument. In one study participants rated the quality of identical pieces of clothing. It was found that the order in which participants evaluated each item of clothing strongly influenced their judgements, yet when asked about their reasons for preferring one item to another, participants did not mention order, and indeed denied that order matter when asked directly. In another study participants indicated that a distracting noise influenced their evaluation of a film, but the data showed that this was not the case. Other participants believed that warnings about the danger of an electric shock influenced their predictions of how much shock they could endure, but results showed that the warnings had no effect.

The reasons for a person’s moral judgements may be equally opaque to them. Haidt et al. (2000) examined participants’ justifications for their moral judgements, and found that for judgements of moral dilemmas thought to rely on intuition, participants were often unable to justify their responses (e.g. “I know it is wrong, but I just can’t come up with the reason why” p.14). Similar findings were reported by Haidt and Wheatley (2005), who hypnotised highly suggestible participants to feel disgust to a particular word. The word was either present or absent in a description of a moral transgression and it was found that when the word was present, participants would rate the transgression as more morally wrong. However when asked about their judgements,

---

10 Unpublished manuscript cited > 90
participants in the suggestion condition often could not justify their answers, or they
gave justifications that made assumptions about information not presented in the
descriptions of the moral transgressions.

Given that people often do not have insight into what influences their judgements it is necessary to experimentally verify that the motives identified as important in Study 1, are in fact important for making goodness judgements of actions.

Motives as indication of character?

It has been proposed that once an action is identified as intentional, evaluations of that action are then made based upon the consequences of the action as well as the reason or causal history explanation for the agent’s action (Alicke, 2000; Guglielmo, Monroe, & Malle, 2009). According to this account, motives then influence goodness judgements of actions directly, depending on the concept people have of a good action and how specific motives fit into that concept, whether they are the “right” or “wrong” kinds of motives for performing good actions.

However, there is an alternative explanation to motives directly influencing goodness judgements of actions. It could be that motives influence goodness judgements of an agent’s actions by altering perceptions of the agent’s character, such that the goodness of an action depends on the goodness of the person performing the action.

Reeder and colleagues identified a link between motive and character judgements, finding that motives influence attributions of traits, and in one study explicitly character judgements (Reeder et al., 2002). A link from character evaluations to moral judgements of actions has also been found (Alicke, 1992, 2000; Nadler & McDonnell, 2011). People who are not liked or judged to be bad, are blamed more for their actions (Alicke, 1992; 2000). In one study Nadler and McDonnell (2011) found that positive or negative evaluations of a person whose actions led to a harmful accident influenced blame and responsibility judgements for the accident.

Pizarro and Tannenbaum (2012) argue that moral evaluations of actions are in essence moral evaluations of a person’s character. The authors point out that findings in the moral judgement and blame literature that are interpreted as processing biases
(Alicke, 2000), can more parsimoniously be explained if we understand judgements of actions as – at least in part – judgements of character. According to their hypothesis people evaluate an action by asking, “Is the sort of thing that a bad (and presumably good) person would do”.

To assess whether evaluations of an agents character can account for the potential influence of the agent’s motives on goodness judgements of actions, character judgements will be evaluated alongside goodness judgements. By including measures of character it can be determined whether any observed change in judgements of moral goodness is due to motive, perception of character, or both.

Study 2

In Study 1 participants identified several motives that are central to what it means for an action to be morally good. Though participants identified these motives as central to the concept of a good action in Study 1, people do not always have insight into the reasons behind their judgements and actions (Nisbett & Wilson, 1977), and moral evaluations (Haidt, 2001). Study 2 will seek to experimentally establish that the motives identified as important by participants, are in fact relevant for goodness judgements of actions.

Participants will read descriptions of people motivated by one of three factors identified in Study 1: other interest (i.e., motivated by the interests of others. In this study this was exemplified by a desire to actively help others or to avoid negatively affecting them); principle (i.e., motivated to act because of a principle, rule, standard or moral belief. In this study agents have a moral belief that something is important); or self interest (i.e., motivated by self interest. In this study agents were motivated by a desire to obviously benefit themselves). Participants are also motivated by a neutral fourth factor intended to be morally neutral. Motives in the neutral condition are characterised by a desire to perform a mundane action that does not positively or negatively influence others.

The other two motives identified in Study 1 were no motive to harm and sympathy. A motive to not harm others was not investigated in this study, because research has demonstrated that having intent to harm others increases how morally bad an action is considered to be (Cushman, 2008). However, given my argument that
judgements of morally good and bad actions are made in different ways, excluding this motive because of findings in the transgression literature was an oversight. Actions motivated by sympathy will be considered in the next chapter.

I hypothesized that the motives identified in Study 1 would influence goodness judgements, and predicted firstly that actions motivated by the interests of others and by principles would be evaluated as more morally good than actions resulting from a neutral motive. I had no prediction about the difference between the other interest and principle conditions. Second, I predicted that goodness judgements of actions motivated by self interest would not differ from the neutral motive. Third, I predicted that actions motivated by other interest and self interest would be judged less morally good than actions motivated by other interest alone. I also examined, but did not make a prediction about, whether an action motivated by a combination of other interest and self interest differs from the neutral condition. Participants in Study 1 indicated that good actions should not be motivated by self interest, but it is not clear whether the presence of self interest will render actions as merely less morally good than when self interest is absent, or not morally good at all (i.e., no different from the neutral conditions). Finally, I examined whether the relationship between motive and goodness judgements is mediated by character judgements for which I did not have a specific prediction.

Method
Participants
198 students (146 females, 52 males (Mage = 19.88, SD = 2.46) studying first or second year psychology at The University of Otago gave informed consent before participating in this study in exchange for partial course credit.

Materials and Procedure
This study was one among several unrelated studies that participants completed individually. Participants were instructed that they would be presented with several situations and asked what they think about the actions of the people in each situation. Stimuli were presented using SuperLab 4 experiment presentation software on iMac 21-inch computers in private experimental cubicles. Participants read vignettes and responded using the keyboard.
Participants read eight vignettes which were designed to have an outcome that is morally neutral (see Table 3.1; full list in Appendix C)\(^{11}\). The stated motive of the agents in each set of eight vignettes was manipulated between subjects. The vignettes were as similar as possible except for the motivation of the agent: in the other interest condition the agent was motivated to positively act in the interest of others or avoid negatively affecting others; in the principle condition agents were motivated to act in accordance with a strong moral belief that something was important; in the self interest condition the agent was motivated by their own interests; in the self/other interest condition agents were motivated both by the concern for their own interests \textit{and} the interests of others (order of information was counterbalanced); in the neutral condition the agent was motivated to act in a way that did not influence others.

This study combines the data from two similar samples collected at different points in time: The neutral, other interest and principle conditions were collected with the first sample, and the self interest and self/other interest conditions were collected with the second.

After reading each vignette participants answered three questions, with each question accompanied by a 9 point scale: \textit{How morally good or bad was it that X did Y?} 1(Extremely morally bad) – 5(Morally neutral) – 9(Extremely morally good), \textit{Do you think that X is mainly a good or bad person?} 1(Mainly a bad person) – 9(Mainly a good person), \textit{Do you think that X has good moral standards?} 1(Not at all) – 9(Completely). Participants indicated their answer by pressing the number key on the keyboard that corresponded with the number on the scale. The latter two questions were used to assess character (taken from Inbar, Pizarro & Cushman, 2011).

\(^{11}\) To be verified by judgements in the neutral condition.
Table 3.1. *Example vignettes for each condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vignette</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other interest</td>
<td>It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he doesn’t want to annoy his neighbours by playing music late at night.</td>
</tr>
<tr>
<td>Principle</td>
<td>It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he has a moral belief that you shouldn’t behave selfishly so he turns off the music and goes to bed so he doesn’t annoy his neighbours by playing music late at night.</td>
</tr>
<tr>
<td>Self interest</td>
<td>It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he doesn’t want to annoy the girl next door who he is trying to impress.</td>
</tr>
<tr>
<td>Self/other interest</td>
<td>It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he doesn’t want to annoy his neighbours by playing music late at night and he also doesn’t want to annoy the girl next door who he is trying to impress.</td>
</tr>
<tr>
<td>Neutral</td>
<td>It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he is tired.</td>
</tr>
</tbody>
</table>

*Note.* In this set of vignettes the actor is performing a morally good action by avoiding negatively affecting others. In other vignettes agents actively performed morally good actions.

**Results**

Participants’ responses to the morality question were averaged across vignettes (*α* = .82) and the two character questions were averaged (*α* = .88) and the mean across vignettes calculated (*α* = .90).

Means for the counterbalanced groups for the self/other interest condition were not significantly different *t*(41) = 1.55, *p* = 0.13, and so the counterbalanced groups were combined for this analysis. A one way between subjects ANOVA comparing the five motive conditions (other interest, principle, self interest, self/other interest, neutral) was significant, *F*(4,193) = 19.95, *p* < .01, *η*² = .28.
Follow up $t$ tests were conducted to compare means specific to the hypotheses\(^{12}\) (see Figure 3.1). The other interest and principle conditions were predicted to increase goodness judgements compared to the neutral condition, and this is what was found. Actions motivated by other interest ($M = 6.85, SD = 1.07$) received significantly higher moral goodness ratings than actions following from a neutral motive ($M = 5.66, SD = .67$), $t(69) = 5.59, p < .01, 95\%$ CI $[-1.61, - .76], d = 1.33$. Actions motivated by principles ($M = 6.58, SD = 0.78$) received significantly higher moral goodness ratings than actions resulting from neutral motives, $t(72) = 5.40, p < .01, 95\%$ CI $[-1.26, - .58], d = 1.26$. Moral goodness judgements for the other interest and principle conditions did not differ, $t(73) = 1.24, p = .22, 95\%$ CI $[-.16, .70], d = .29$.

As predicted, there was no significant difference between self interest ($M = 5.46, SD = 0.67$) and the neutral condition, $t(78) = 1.36, p = .18, 95\%$ CI $[-0.09, 0.51], d = 0.31$. Also as predicted, goodness judgements were significantly lower for the combined self/other interest condition ($M = 6.37, SD = 1.05$) compared with the other interest condition, $t(77) = 2.00, p = .049, 95\%$ CI $[0.00, 0.95], d = 0.45$. The self/other interest condition also received significantly higher goodness judgements than the neutral condition, $t(76) = 3.460, p < .01, 95\%$ CI $[-1.116, -0.301], d = 0.806$, see Figure 3.1.

\(^{12}\) I didn’t control for familywise error rate; I don’t find the arguments for adjusting in this context compelling. Furthermore I don’t have the power to be conservative for example observed power to detect a medium effect $d = .5, n = 69$, is $.53$ at $p = .05$ with the Bonferroni corrected alpha/6 = .008, power to detect a medium effect drops to .258. From 50% to 25%.
Character judgements

The same analyses were conducted for character judgements. A significant difference was found between the other interest ($M = 7.10$, $SD = 1.00$) and neutral conditions ($M = 6.18$, $SD = 0.92$), $t(69) = 4.07$, $p < .01$, 95% CI [-1.39, -0.47], as well as between principle ($M = 7.11$, $SD = 0.92$) and neutral conditions, $t(72) = 4.37$, $p < .01$, 95% CI [-1.36, -0.51]. There was also no significant difference between other interest and principle conditions $t(73) = 0.01$, 95% CI [-0.44, 0.44].

The self interest condition ($M = 5.91$, $SD = 0.89$) did not differ significantly from the neutral condition, $t(78) = 1.27$, $p = .21$, 95% CI [-0.15, 0.66], and the self/other interest condition ($M = 6.58$, $SD = 1.17$) had significantly lower character scores than the other interest condition, $t(77) = 2.11$, $p = .04$, 95% CI [0.03, 1.01]. There was no significant difference between the self/other interest condition and the neutral condition, $t(76) = 1.68$, $p = .10$, 95% CI [-0.89, 0.07].
Mediation

To test whether character judgements mediated the relationship between motives and goodness judgements I conducted a series of mediation analyses on the pairs of means identified to differ significantly on moral goodness judgements. In each case regression equations were calculated for each path and a bootstrap mediation analyses (Preacher & Hayes, 2004) was conducted (10000 bootstrap samples) to estimate the size of the indirect path (paths $a \cdot b$, Figure 3.2). Path $c'$ is the effect of motivation on moral judgement while controlling for mediation by character. Table 3.2 displays path coefficients and mediation significance for each pairwise comparison.

Model 1.

![Model 1](image)

Model 2.

![Model 2](image)

Figure 3.2. Model 1. Overall effect of an agent’s motive on an observer’s goodness judgements. Model 2 shows the motive effect on moral judgement being mediated by character.

When the type of motive (other interest or neutral) was used to predict goodness judgements, character judgements were shown to significantly mediate the relationship. The effect of motive condition on goodness judgements was still significant when controlling for character. Similarly when the type of motive (principle or neutral) was used to predict goodness judgements, character judgement significantly mediated the relationship and the direct effect of reason condition on goodness judgements was still significant.

It was shown that moral judgement scores for the self/other interest condition were significantly lower than in the other interest condition. This relationship appears to be mediated by character judgements, and when controlling for character judgements, motive type (self/other interest or other interest) no longer significantly predicts
goodness judgements. When motive type (self/other interest or neutral) predicted goodness judgements, the relationship was not significantly mediated by character judgements.

Table 3.2. Path coefficients and significance for each pairwise comparison

<table>
<thead>
<tr>
<th></th>
<th>a B</th>
<th>SE</th>
<th>b B</th>
<th>SE</th>
<th>c B</th>
<th>SE</th>
<th>c' B</th>
<th>SE</th>
<th>ab B</th>
<th>SE</th>
<th>95%LB</th>
<th>95%UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other interest–Neutral</td>
<td>.74*</td>
<td>.07</td>
<td>.93*</td>
<td>.23</td>
<td>1.19*</td>
<td>.21</td>
<td>.50*</td>
<td>.14</td>
<td>.69*</td>
<td>.70</td>
<td>.33</td>
<td>1.11</td>
</tr>
<tr>
<td>Principle – Neutral</td>
<td>.58*</td>
<td>.06</td>
<td>.47*</td>
<td>.11</td>
<td>.46*</td>
<td>.09</td>
<td>.19*</td>
<td>.07</td>
<td>.27*</td>
<td>.08</td>
<td>.13</td>
<td>.44</td>
</tr>
<tr>
<td>Self/other interest – Other interest</td>
<td>.03*</td>
<td>.01</td>
<td>.86*</td>
<td>.05</td>
<td>-.02*</td>
<td>.01</td>
<td>.00</td>
<td>.01</td>
<td>.02*</td>
<td>.01</td>
<td>-.04</td>
<td>-.00</td>
</tr>
<tr>
<td>Self/other interest – Neutral</td>
<td>.02</td>
<td>.01</td>
<td>.67*</td>
<td>.06</td>
<td>.03*</td>
<td>.01</td>
<td>.02*</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>-.00</td>
<td>.03</td>
</tr>
</tbody>
</table>

* p < .05

Discussion

As predicted, motives of other interest and principle increased how morally good an action was perceived to be relative to the neutral condition. It was also found that when self-interest was paired with other interest, actions were judged as less morally good than when other interest was the only motive. Furthermore the self/other interest condition was judged more morally good than the neutral condition. These results provide experimental evidence for the influence of the motives identified in Study 1 on goodness judgements of actions.

The findings of this study, and of Study 1, provide some of the first insights into which specific motives are part of the concept of a morally good action, and how these motives influence goodness judgements of actions. The findings are consistent with previous research indicating that motives influence goodness judgements (Baird & Astinton, 2004; Janoff-Bulman et al., 2009), but go beyond simply identifying that motives matter, to show the importance of specific motives.
Furthermore, the results suggest that the influence of motive on goodness judgements is partially mediated by character judgements, which is consistent with the idea that the agent’s actions were judged morally good because of how motives reflect on the agents themselves. Character judgements mediated the relationship between all motive combinations and goodness judgements, with the exception of the difference between self/other interest and neutral were used to predict goodness judgements. Previous research has identified links between motive and character (Reeder et al., 2002), and character and moral judgements of actions (Nadler & McDonnell, 2011) and recently Pizarro and Tannenbaum (2012) have argued that factors which influence moral judgements of actions, such as intentionality and control, influence judgements of actions because they indicate things about why a person is performing an action, which is revealing of the person’s character. Motive is potentially more revealing of character than even the outcomes of people’s actions, so the finding that motive influences goodness judgements of actions is consistent with the account that actions are good because they are the kinds of actions that a good person would perform. Though previous research has examined the relationship between judgements of actions and character, to the best of my knowledge, this study is the first to document mediation of the relationship between motive and goodness judgements by character.

Other interest

Being motivated by the interest of others increased how morally good an action was judged to be. This effect was partially mediated by character judgements.

It is important to note that the other interest condition did not distinguish between two senses of “other interest”; a desire to benefit others (e.g., “Daniel decided to work late because he wanted to help his work colleagues by getting his work done”), and a desire to not negatively affect others (e.g., “Emma is sick and decides not to go to work today because she doesn’t want to make her work colleagues sick”). These two ways of being motivated by the interest of others may not be equivalent. Janoff-Bulman et al., (2009) argued that judgements of someone actively performing an action and inhibiting an action are influenced by different motivational systems. However in this study responses were nearly identical for both actively helping ($M = 6.87$, $SD = 1.07$), and acting to avoid negatively affecting others ($M = 6.81$, $SD = 1.26$), $t(35) = 0.37$, $p$
= .71, 95% CI[-0.34, 0.24]. These two ways of acting were also used in the principle condition, again without significant differences between performing a positive action ($M = 6.61, SD = 0.85$) and acting to avoid a negative outcome ($M = 6.53, SD = 0.92$), $t(38) = 0.60, p = .55$, 95% CI[-0.35, 0.19].

**Principle**

Being motivated by a principle increased how morally good an action was judged to be. This effect was partially mediated by character judgements.

Agents were described as being motivated to act consistently with a principle (e.g., that the environment is important, to not harm others, that working hard is important, and to help others); however, it is likely that principles influenced goodness judgements because of the content of the particular principles not just that an agent was motivated by a principle. If the principles were not ones participants shared, participant’s goodness judgements may not have been affected, as people are intolerant of others who do not share their moral beliefs (Skitka et al., 2005). However, though an observer may not share a principle with an agent (and the observer’s goodness judgements of the agent’s actions may not be as great as if the observer shared the principle), it may still be that by acting in a principled way, the presence of a principle increases the observer’s goodness judgements relative to a situation where the agent performs the same action without a principled belief (or another good motive, i.e. other interest).

In this study the content of the principles were in some instances similar to the other interest condition. For example, in the other interest condition, “Emma is sick and decides not to go to work today because she doesn’t want to make her work colleagues sick”, and in the principle condition “Emma is sick and decides not to go to work today because she has a moral belief that not harming other people is important and she doesn’t want to make her work colleagues sick”. To clearly distinguish motivation by a principle from other interest, the principles used should have been those that did not overlap with other interest (e.g., telling the truth or working hard\textsuperscript{13}).

\textsuperscript{13} Working hard is an example of a principle that was used in the study that did not overlap with other interest.
**Self interest**

As predicted, actions motivated by self interest received similar goodness judgements compared with actions in the neutral condition. Also as predicted, when a morally good action is also motivated by self interest (in this case when an action was motivated by other interest and self interest), the action is judged to be less morally good. As with the other motives, the influence of self interest on goodness judgements was partially mediated by character judgements.

A number of participants in Study 1 suggested that for an action to be morally good, people should not benefit from, or be motivated to benefit from, their actions. The results of Study 2, however, are inconsistent with this suggestion, as actions that were motivated by other interest and self interest, were still evaluated to be more morally good than the neutral condition. However it is likely that the magnitude of the intended benefit of a selfish motive will influence judgements. Performing a good action because a person wants to help others, but also desires to benefit a little from their actions, seems likely to have a smaller influence on goodness judgements than if they desire to benefit greatly. For behaviour explanations, when a person has more than one motive for their actions, an observer must somehow work out which motive was likely the most important reason for the agent’s actions. For example, if someone is motivated to work extra hours on a work project to reduce the workload of their overworked colleagues, and also thinks that either working extra hours will slightly increase their long-term chances of a promotion, or will have a large influence on an imminent promotion, it seems likely that being motivated by the immediate large reward of a promotion will influence observers’ interpretation of self interest as being the motive more responsible for the agent’s actions, rather than if the person was motivated by the distant possibility of a promotion.

**Neutral condition**

The neutral condition was intended to represent a non-moral action that the other motives could be compared to. However, the actions in the neutral condition could be characterised as self interested. Indeed, all non-moral actions are, in a sense, self interested, in the sense that even the most mundane actions, such as choosing what to eat for lunch or turning on the TV, are motivated by the interests of the person...
performing them. The practical distinction between the neutral condition and self interest made in this study, then, may be quantitative rather than qualitative, such that the tangible benefits to the actor in the latter case is more obvious.\footnote{The neutral motive and self interest were run at separate times. When the neutral condition was run the only real concern was to use a motive for comparison that was not obviously one of the other motives or morally valenced.}

It was intended that the neutral condition was considered morally neutral; however, the neutral condition was rated significantly more morally good than the middle point of the scale (which was labelled morally neutral). Any mean will deviate from the midpoint with enough power to detect an effect, and it is not obvious whether a small deviation from the neutral point should still be considered morally neutral, or any deviation from the middle point should be treated as a good (or bad) action. Whether participants’ responses to the neutral condition should be thought of as neutral or slightly morally good is not really a problem for the study, as the hypothesis is that the right kinds of motives increase how morally good an action is evaluated to be. The study was designed to be able to argue that the right kind of motive is sufficient to make a neutral action morally good, but a more conservative interpretation of the data is that an action performed for the right kind of reason increases how morally good the action is evaluated to be.\footnote{Though the data shows that three vignettes in the neutral condition do not differ significantly from the midpoint, so, using only those 3 vignettes, an argument could be made that the right kind of motive was sufficient to make a neutral action morally good.}

A related issue is the use of a continuous scale, from 1 (extremely morally bad), 5 (morally neutral), to 9 (extremely morally good). I have argued that morality is not a continuum and yet in this study I assess morality using a scale from morally bad to good. A better way of assessing good and bad actions at the same time would be to use two separate scales, one asking how morally good an action is, and the other assessing moral wrongdoing.

Conclusion
This study builds on the work of Study 1 and provides experimental evidence that the motives self interest and principle increase how morally good an action is evaluated to be, while self interest decreases goodness judgements. The results are consistent with previous research indicating the importance of motives for goodness judgements of actions (Baird & Astington, 2004; Janoff-Bulman et al., 2009), and extend these
findings by identifying the specific motives associated with a morally good action. It was also found that the relationship between motives and goodness judgements was partially mediated by character judgements. This finding is consistent with previous research reporting evidence for an association between motives and character judgements (Reeder et al., 2002), and character judgements and moral judgements (Alicke, 2000; Nadler & McDonnell, 2011).

One motivation not examined in this study, but identified in Study 1 as important to the concept of a morally good action is sympathy. In the next study I examine the influence of sympathy on goodness judgements of actions. Sympathy was examined separately for practical reasons, as part of a related reinterpretation of an existing finding in the literature (which will be discussed in Chapter 5). However, as the next chapter will reveal, there are other important reasons to distinguish sympathy from other motivations.
Chapter 4

Sympathy and moral goodness

In this chapter I continue my examination of the right kinds of reasons for performing morally good actions. Sympathy was identified in Study 1 as a right kind of reason for performing a morally good action. In this study the influence of sympathy on goodness judgements of actions is experimentally examined. The role of character judgements as a mediator between reasons and goodness judgements is also considered.

In this chapter I examine the influence of sympathy on goodness judgements of actions. Study 2 experimentally examined the influence of several motives identified in Study 1. Sympathy is considered separately from the motives examined in Study 2 as it has particular relevance to a reinterpretation of existing findings in the literature, which will be examined in chapter 5. However there are other important reasons to distinguish sympathy from the other motives.

Sympathy

In Study 1 participants mentioned that the moral good is associated with people feeling empathy and compassion. Sympathy, empathy and compassion are terms that have been used interchangeably in the literature. Empathy has been characterized as both the experience of an emotion that is similar to the emotion that the person they are empathizing with is feeling (Eisenberg, 2000; Hoffman, 2000) and as an emotional response that involves sympathy, compassion, tenderness, and warmth (Batson, Early, & Salvarni, 1997; Batson, Eklund, Chernok, Hoyt, & Ortiz, 2007). The latter description of empathy is much closer to what many other authors refer to as sympathy, which is an emotion involving concern for others (Tangney, Stuewig & Mashek, 2007). More recently Goetz, Keltner and Simon-Thomas (2010) have labelled a feeling resulting from witnessing the suffering of others that motivates a subsequent desire to help, as compassion. Goetz et al. (2010) argue that compassion is very similar to how others describe sympathy, but that they prefer the former term as it covers a broader set of states. Goetz et al. note that lexical studies of emotion terms often find that
compassion and sympathy are grouped together. So in common usage and in technical descriptions these concepts are closely associated with each other. It has been found that a consequence of feeling sympathy is an increase in prosocial motivation (Batson et al, 1997), and helping behaviour (Goetz et al., 2010). Given that differences between sympathy, empathy and compassion are not clear in the technical literature, or common usage, I have chosen to use the term sympathy, as it is consistent with the terminology used in research relevant to work in chapter 5.

Sympathy differs from the motives identified in Study 1. Unlike the other motives that are used in reason explanations of behaviour, sympathy is often used as a causal history explanation. Unlike the motives other interest, principle and self interest, sympathy is not a belief or desire itself, but rather an emotion that influences beliefs and desires, and so it is a cause for why a person has a belief or desires to act in a particular way. For example, saying that her motive (or motivation) for calling her friend an idiot is that she is an angry sort of person, does not make sense; however, saying the reason she called her friend an idiot is that she is an angry sort of person, fits with common usage. Motive and motivation do not work as descriptions when an explanation of behaviour is not something that the agent is likely to be thinking about at the time of an intentional action. And so I will avoid referring to sympathy as a motive, but will talk about it as a reason for performing an action, as in, the reason she paid for his lunch is that she felt sympathy for him\textsuperscript{16}.

Another distinction between motives and sympathy is that motives explain why a person performed a good action, while sympathy explains why a person performed an action and how they were feeling when they decided to act. Why a good action was performed has been shown to be important for goodness judgements, but it may also matter whether a person felt the right kind of way when deciding to perform, or while performing a good action. Supporting this idea Ames, Flynn and Weber (2004) found that when a helper was motivated by positive affect and not their role, obligation, or calculation of costs and benefits, the person being helped liked the helper more, and was more likely to reciprocate.

\textsuperscript{16} This is not to be confused with the term “reason explanation” which is a type of behaviour explanation for intentional actions that makes use of beliefs and desires.
Study 3
In this study I examine how the presence of sympathy as a reason for acting influences observers’ judgements of the moral goodness of an agent’s actions. Based upon the findings of Studies 1 and 2, I hypothesise that sympathy as a reason for performing an action will increase how morally good that action is evaluated to be, and predict that participants will rate the actions of an agent who performs a good action out of sympathy as more morally good than an agent who performs a good action for a non-sympathy reason.

In Study 2 the influence of motive was examined using vignettes with neutral outcomes. In this study agents perform actions with outcomes that are morally good, even without explicit mention of motives or reasons. The claim for both studies is that motive and reasons influence goodness judgements to make them more or less morally good, regardless of whether the action performed would be considered good or neutral without further information about motives.17

Method
Participants
One hundred and fourteen18 first and second year psychology students at Otago University gave informed consent before participating in return for course credit.

Materials and Procedure
This study was one among several unrelated studies. Stimuli were presented using SuperLab 4 experiment presentation software on iMac 21-inch computers in private experimental cubicles. To assess the influence of sympathy participants read vignettes and responded using the keyboard.

Participants read vignettes describing a person performing morally good actions by performing an action that benefited others. In the sympathy conditions the actors explicitly cite sympathy as a reason for their actions, and also report that their actions are “the right thing to do”; in the control condition they only state that their actions were

17 The design with the neutral outcome turns out to be better for demonstrating the influence of motives. However it is necessary for my argument that motives and reasons influence actions with morally good outcomes.
18 Demographic data were not recorded.
the right thing to do (identified in Study 1 as a synonym for saying that something is morally good).

Although actors in the control condition do not explicitly say they feel sympathy, it is possible that participants will spontaneously infer that they feel sympathy on the basis of their helpful actions (Hassin, Aarts, & Ferguson, 2005). To control and test for such inferences, the question posed to the actor in the vignette was also varied. In the volunteered conditions, actors were asked why they performed the action, to which they reply in the sympathy condition that “I felt enormous sympathy and it was the right thing to do”, and in the control condition, “It was the right thing to do”. In the questioned about sympathy conditions, actors are explicitly asked whether they performed the action because they felt sympathy, and they reply either that they “Felt enormous sympathy … and it was the right thing to do”, or, “Not really, I just think it’s the right thing to do”. Examples of vignettes appear in Table 4.1, and the full set appears in Appendix D.

Participants were instructed that they would be presented with several situations, and asked what they thought about the people’s actions in each situation, and provided with an example. Participants then read and evaluated three sympathy or no-sympathy vignettes from the volunteered condition (the presentation order of the three vignettes was randomized). The content of the vignettes was identical for the sympathy and no-sympathy conditions, except for the presence or absence of sympathy. Participants were then instructed that they would read about the same situations but that the reasons for the actions of the people in the vignettes had changed. Participants then read and evaluated vignettes from the other sympathy condition. Whether participants saw the sympathy or no-sympathy condition first was randomized between participants.

After each vignette participants were asked three questions. Each question was accompanied by a seven-point scale. Participants indicated their answer by pressing the number key on the keyboard that corresponded with the number on the scale. *How praiseworthy was it that X did Y? 1(not at all praiseworthy) – 7(Extremely praiseworthy), How morally good was it that X did Y? 1(not at all morally good) – 7(Extremely morally good), How positively should X be evaluated for Y? 1(not at all positively) – 7(Extremely positively).*

After completing a filler task (part of an unrelated study that took approximately
30 minutes), participants were then instructed that we were again interested in their evaluations of people’s actions. They then completed sympathy and no-sympathy conditions for *questioned* vignettes (the content of which differed from those in the *volunteered* vignettes), in the same manner as for *volunteered* vignettes.

### Table 4.1. Example vignettes for volunteered and questioned vignette types, and sympathy and no sympathy conditions

<table>
<thead>
<tr>
<th>Vignette type</th>
<th>Sympathy</th>
<th>No Sympathy</th>
</tr>
</thead>
</table>
| **Volunteered** | Jackie donated $200 to help the victims of a flood. When asked why she donated the money, she replied, “I felt enormous sympathy for the victims and it was the right thing to do”.
Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “I feel enormous sympathy for him and I think it’s the right thing to do”.
Mathew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Mathew goes to visit his friend in the hospital. Mathew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “Not really, I just think it’s the right thing to do”.
| **Questioned** | Jackie donated $200 to help the victims of a flood. When asked why she donated the money she replied, “It was the right thing to do”.
Mathew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Mathew goes to visit his friend in the hospital. Mathew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “Not really, I just think it’s the right thing to do”.

### Results

Scores for the three questions following each vignette were averaged to form a moral judgement score ($\alpha = .87$), and mean moral judgement scores were then calculated across the 3 vignettes ($\alpha = .75$). They appear in Figure 4.1. A 2 (sympathy: present, absent; within) x 2 (vignette type: volunteered versus questioned; within) x 2 (sympathy order: sympathy first, no sympathy first; between) ANOVA was conducted. The order main effect was significant $F(1,111) = 4.14, p = .04$, $\eta^2_p = .04$; moral judgements ratings

---

19 Agents were described as being asked about their reasons for acting. A benefit of this design is that almost exclusively in the literature – and in Study 2 – information about mental states and reasons for acting are given by an omniscient observer. People do not have access to motive and reason information directly, and the only way they get it is to infer it themselves or find out from the agent or from a third party. Phrasing the study in this way is a more realistic way of providing reason information.
were higher when the no sympathy condition was presented first (no sympathy first, $M = 5.92, SD = 0.80$; sympathy first $M = 5.73, SD = 0.81$), but none of the interactions with order were significant\textsuperscript{20}.

There were also main effects of vignette type and sympathy. \textit{Volunteered} vignettes received significantly higher moral goodness ratings ($M = 6.2, SD = 0.56$) than \textit{questioned} vignettes ($M = 5.46, SD = 0.80$), $F(1,109) = 136.87, p < .01, \eta^2_p = .56$; and vignettes including sympathy ($M = 6.06, SD = 0.73$) received higher ratings than vignettes without sympathy ($M = 5.58, SD = 0.73$), $F(1,109) = 53.97, p < .01 \eta^2_p = .33$. The interaction between vignette type and sympathy conditions was significant, $F(1,109) = 49.23, p < .01, \eta^2_p = .31$.

Paired samples $t$ tests found that for \textit{volunteered} vignettes there was no significant difference between the sympathy ($M = 6.23, SD = 0.60$) and no sympathy condition ($M = 6.21, SD = 0.59$), $t(112) = 0.51, p = .61, 95\% \text{ CI } [-0.07, 0.12], d = 0.03$. For \textit{questioned} vignettes, moral goodness ratings in the sympathy condition ($M = 5.92, SD = 0.96$) were significantly higher than the no sympathy condition ($M = 4.97, SD = 1.30$), $t(113) = 7.75, p < .01, 95\% \text{ CI } [0.71, 1.19], d = 0.81$.

\textsuperscript{20} The order interactions were not significant order x sympathy x vignette $F(1,111) = 2.125, p = .148, \eta^2_p = .019$, order x sympathy, $F(1,111) = 2.804, p = .097, \eta^2_p = .025$, order x vignette, $F(1,111) = 0.727, p = .396, \eta^2_p = .007$. 

53
Discussion

The results for questioned vignettes support the hypothesis that sympathy as a reason increases how morally good an action is evaluated to be. This is consistent with Study 2 which found that an agent’s reasons for performing actions influenced goodness judgements. The results for the volunteered vignettes, however, do not support the hypothesis as volunteered vignettes did not show a significant difference between sympathy conditions.

Questioned vignettes were included in this study because it is possible that in the no sympathy condition for volunteered vignettes participants would infer sympathy from the actions agents performed to help others in need. Inferring reasons is central to behaviour explanations (Guglielmo, Monroe, & Malle, 2009; Malle, 2004) and there is evidence that inferring reasons is automatic (Hassin, Aarts, & Ferguson, 2005; Overwalle, Duynslaeger, Coomans, & Timmermans, 2012). By including the questioned vignettes and explicitly stating the agent did not act because of sympathy, it was hoped we could control participants’ assumptions about sympathy. Since sympathy influenced goodness judgements for questioned and not volunteered vignettes, the results suggest that participants did indeed infer sympathy from the agent’s action in the

Figure 4.1. Mean moral goodness scores (+SE) for sympathy and no sympathy condition for volunteered and questioned vignettes.
no sympathy condition for volunteered vignettes.

However, there are several other explanations for why volunteered sympathy did not increase goodness judgements. It is possible that an effect of sympathy was not detected for volunteered vignettes because of a ceiling effect. Means for both conditions (sympathy, $M = 6.23$, $SD = 0.60$, no-sympathy, $M = 6.21$, $SD = 0.59$) were at the high end of the seven-point scale. It is also possible that the study lacked adequate power to detect an effect. A post hoc power estimate indicates that power was adequate to detect a medium effect, $\eta_p^2 = .06$, power = .91), but is underpowered to detect a small effect size, $\eta_p^2 = .01$, power = 0.26.

Another explanation for why goodness judgements were influence by sympathy for questioned vignettes and not volunteered vignettes is that denying feeling sympathy could decrease goodness judgements because the character of a person who denies feeling sympathy for the suffering of others is evaluated less positively than one who affirms that they feel sympathy, or who simply says nothing. This diminished evaluation of a person’s character could then decrease how morally good their actions are judged to be, as there is evidence that character judgements can influence judgements of actions (Study 2; Alicke, 1992, 2000).

Similarly denying feeling sympathy could also decrease goodness judgements not because of how it reflects on a person’s character but because part of what makes an action good is that a person has the right kind of reason for performing the action, and if a person denies being motivated by a right kind of reason, the action may not be judged as morally good.

It could also be argued that the results are partially due to methodological artefacts. For example, the volunteered vignettes were always presented before questioned vignettes. Second, although in each vignette the agent performed an action that benefited other people, the contents of each kind of vignettes differed, and so the effect of sympathy for the questioned vignettes could be due to the particular vignettes used.

Third, the sympathy condition for both types of vignettes there were two motives (sympathy and the agent thought it was the right thing to do). In the no sympathy condition there is only one motive (it was the right thing to do). This was done to try to keep the information consistent between vignettes except for the presence
of sympathy information. This design is open to the criticism that for questioned vignettes it is not sympathy that is influencing goodness judgements but the fact that in the sympathy condition there are two reasons, while in the no sympathy condition there is only one reason. This criticism is only compelling, however, if any two reasons increase how morally good the action is judged to be. My hypothesis is that only certain reasons for acting increase how morally good an action is judged to be, which is supported by Study 1. If both reasons are the right kind of reasons, then higher moral goodness scores in the two reasons condition would still indicate that sympathy is a right kind of reason.\footnote{There are difficulties with all designs but sympathy versus no motive was not used, because comparing a condition with a motive versus no motives left participants assumptions about motives completely open. Sympathy vs it was the right thing to do was not used because that design just allows for a direct comparison between two motives sympathy and the right thing to do.}

Given the number of plausible explanations for the data, two additional studies, 4a and 4b, were conducted to test two of the more plausible explanations for why sympathy influenced goodness judgements for questioned but not volunteered vignettes. Study 4a examines if a ceiling effect can account for the fact that sympathy did not influence goodness judgements for volunteered vignettes. And study 4b examines whether denying feeling sympathy reduces goodness judgements because of negative evaluation of a person’s character.

\textit{Study 4a}

In Study 3 sympathy influenced goodness judgements for questioned and not volunteered vignettes, which suggests that for the no-sympathy version of volunteered vignettes participants spontaneously inferred that the agent was motivated by sympathy. However, several alternative interpretations were proposed, including the hypothesis that the outcomes of volunteered vignettes were too good to detect the influence of sympathy on goodness judgements.

To test if a ceiling effect can account for the similar moral goodness ratings for sympathy and no sympathy conditions one of the vignettes was replaced with a vignette with a less morally good outcome, and the extremity of the outcomes for the other two vignettes was reduced (e.g., in Study 3 the agent donated \$200 to charity, in this study \$10 is donated; see Table 4.4). If a ceiling effect can account for the similar goodness
judgments for sympathy and no sympathy volunteered vignettes in Study 3, by reducing the moral goodness of the actions portrayed in the vignettes we should observe a difference in goodness ratings between sympathy conditions.

Method
Participants
One hundred and thirty three students (102 females, 31 males (Mage = 21.80, SD = 7.61) studying first or second year psychology at the University of Otago gave informed consent before participating in this study in exchange for partial course credit.

Procedure and Materials
The procedure and materials were the same as the volunteered vignettes procedure in Study 3 except that the vignettes were changed to reduce how morally good the actions of the agents were by replacing one vignette, and reducing the magnitude of the help given (Table 4.4. and Appendix E). Another difference to Study 3 was that only the first moral judgement question from Study 3 was used22 (How morally good was it that X did Y? 1(not at all morally good) – 7(Extremely morally good)).

Results
Moral goodness scores were calculated for each participant by averaging across vignettes (α = .85). A 2 (sympathy, no sympathy; within) x 2 (order of sympathy conditions: between) ANOVA was conducted. The order main effect was not significant, $F(1,132) = 1.65, p = .20, \eta_p^2 = .01$, however the mean judgements of moral goodness were significantly higher for the sympathy condition ($M = 5.49, SD = 1.14$) than the no sympathy condition ($M = 5.25, SD = 1.20$), $F(1,132) = 9.25, p < 0.01, \eta_p^2 = .07$. The interaction between sympathy and order was not significant, $F = (1,132) = 22$

---

22 Instead of asking how morally good, how much praise and how positive the agent should be evaluated, participants are only asked how morally good was it that the agent did X. The three question morality scale was chosen to be consistent with Pizarro et al (2003). Beyond consistency I do not have reason to favour the three question scale over a the single question. In fact there is evidence that questions about wrongness and blame (and potentially questions of rightness/goodness and praise) address importantly different things, for example an action may not be morally wrong but can still be blameworthy (Cushman, 2008).
2.31, \( p = 0.13, \eta^2_p = .02 \).

**Discussion**

For volunteered vignettes a significant difference was observed between the sympathy and no sympathy conditions; this is consistent with the hypothesis that sympathy as a reason for acting will increase how morally good an agent’s actions are evaluated to be. That a significant difference was identified between sympathy and no sympathy conditions in this study and not Study 3 gives support to the hypothesis that no difference was found in Study 3 because of a ceiling effect. Mean moral judgement scores in Study 3 for voluntary vignettes were approximately \( M = 6.2 \) for both conditions\(^{23}\), and \( M = 5.25 \) in this study for the no sympathy group and \( M = 5.49 \) for the sympathy group. Given these findings it is less likely that sympathy did not influence goodness judgements for *volunteered* vignettes in Study 3 because participants attributed sympathy to the agent in the no sympathy condition.

**Study 4b**

For *questioned* vignettes in Study 3 actions were judged to be more morally good in the sympathy condition than the no sympathy condition. Although I interpret this difference as evidence that the presence of sympathy increases the goodness of an action, an alternative account is that denying sympathy is sufficient to reduce the morality of an action, because such denial reflects poorly on the agent’s character. Previous research (Alicke 1992, 2000; Nadler & McDonnell, 2011) and Study 2 have found that perceptions of a person’s character can influence moral evaluations of their actions.

The hypothesis that for *questioned* vignettes less positive character evaluations can account for the observed difference between sympathy conditions, is examined in Study 4b by measuring character judgements. If the difference in sympathy conditions was due to the influence of character judgements alone, these judgements should fully mediate the relation between sympathy and goodness judgements. Partial mediation, as found between motive and goodness judgements in Study 2, would indicate that, though

\(^{23}\) Study 3 used three questions to assess morality and study 4a used only 1. If you compare the same morality question in both studies you get the same mean for study 4a. Indicating that the change in means is not due to a change in the way morality was assessed.
character partially mediates the relationship, sympathy has a direct effect on goodness judgments.

Method

Participants
One hundred and twenty nine students (91 females, 38 males \(M_{age} = 20.25, SD = 3.67\)) studying first or second year psychology at the University of Otago gave informed consent before participating in this study in exchange for partial course credit.

Procedure and Materials
The procedure was similar to the questioned vignettes procedure used in Study 3 with the exception that only one morality question was asked, and questions about character were also asked (Appendix F). After each vignette participants were asked: How morally good was it that X did Y? 1(not at all morally good) – 7(Extremely morally good). Do you think that X is mainly a good or bad person? 1(Mainly a bad person) – 9(Mainly a good person), Do you think that X has good moral standards? 1(Not at all) – 9(Completely).

Results
Mean moral goodness scores were calculated by averaging across vignettes \(\alpha = .86\). Character scores were calculated by averaging the answer to the two character questions \(\alpha = .83\). Character scores were then averaged across vignettes \(\alpha = .88\). A 2 (sympathy, no sympathy; within) x 2 (order) ANOVA was conducted on mean goodness scores. Moral goodness scores were higher on average for the sympathy condition \(M = 5.78, SD = 0.92\) than the no sympathy condition \(M = 4.73, SD = 1.29\), \(F(1,128) = 79.00, p < .01, \eta^2_p = .38\). The order main effect was not significant \(F(1,128) = 78.89, p < .01, \eta^2_p = .38\), however the interaction between order and sympathy conditions was significant, \(F(1,128) = 5.46, p = .02, \eta^2_p = .04\).

The significant interaction still allows for a simple interpretation of the sympathy main effect, as there was a significant difference between sympathy and no sympathy conditions for both orders (order 1: sympathy \(M = 5.58, SD = 0.89\), no
sympathy ($M = 5.01, SD = 1.01$), $t(65) = 6.62, p < .001, d = 0.82$; order 2: sympathy ($M = 5.77, SD = 0.95$), no sympathy ($M = 4.43, SD = 1.47$), $t(63) = 6.39, p < .01, d = 1.08$.

A 2 (sympathy, no sympathy; within) x 2 (order) ANOVA was also conducted on mean character scores. Moral goodness scores were higher on average for the sympathy condition ($M = 5.98, SD = 0.72$) than the no sympathy condition ($M = 5.57, SD = 0.90$), $F(1,128) = 21.31, p < .01, \eta^2_p = .14$. The order main effect was not significant $F(1,128) = 0.71, p = .40, \eta^2_p < .01$, the interaction between order and sympathy conditions was also not significant, $F(1,128) = 3.03, p = .08, \eta^2_p = .02$.

To test whether character judgements mediated the relationship between sympathy and judgements of moral goodness a bootstrap mediation analysis (Preacher & Hayes, 2004) was conducted (10000 bootstrap samples) to estimate the size of character mediation (paths $a \cdot b$; see Figure 3.2). As seen in Table 4.2, the analysis indicates that character judgements significantly mediate the relationship between sympathy and judgements of moral goodness, but that the direct effect of sympathy (controlling for character; path $c'$) remains significant.

### Table 4.2. Mediation analysis between sympathy condition and goodness judgements by character judgement

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$c'$</th>
<th>$ab$</th>
<th>$95%\ LB$</th>
<th>$95%\ UB$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>B</td>
<td>SE</td>
<td>B</td>
<td>SE</td>
<td></td>
</tr>
<tr>
<td>Sympathy as a predictor</td>
<td>.28*</td>
<td>.10</td>
<td>.60*</td>
<td>.08</td>
<td>.98*</td>
<td>.14</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>.82*</td>
<td>.13</td>
<td>.17*</td>
<td>.07</td>
<td>.05</td>
<td>.31</td>
<td></td>
</tr>
</tbody>
</table>

Note. * $p < .05$. For model see Figure 3.2.

**Discussion**

This study replicated the findings of Study 3 that for questioned vignettes actions in the sympathy condition are rated as more morally good than action in the no sympathy condition. Furthermore it was found that the relationship between sympathy condition and goodness judgements was partially mediated by character judgements, which is not consistent with a hypothesis that less positive evaluations of character when an agent denies acting out of sympathy can entirely account for the difference observed between
sympathy conditions.

Though a significant effect of sympathy condition was found for both vignette types in Studies 4a and 4b, the difference between sympathy and no sympathy conditions for *questioned* vignettes ($\eta_p^2 = .38$) is substantially larger than for *volunteered* vignettes ($\eta_p^2 = .07$). Perhaps this difference in effect size exists because, as discussed, sympathy as a reason for performing a good action increases goodness judgements, and *also* denying acting out of sympathy decreases goodness judgements. It could also be the case that by explicitly asking about sympathy in the *questioned* vignettes, the role of sympathy for the agent’s actions are made more salient to the participant, and consequently sympathy has a greater influence on goodness judgements.

**Study 5a**
Together Studies 3 and 4 provide evidence that the presence of sympathy when performing an action increases how morally good an observer evaluates that action to be, and that this effect is not due to character inferences about actors who deny feeling sympathy. However, it may be the case that another inference about sympathy is responsible for participants’ judgements.

Sympathy is an emotion involving concern for the welfare of others and the alleviation of others’ suffering (Haidt, 2003a), and as a consequence of feeling sympathy a person is motivated to help others (Goetz et al., 2010). Studies 1 and 2 identified that a motive to act in the interest of others is central to the concept of a morally good action, and Study 2 found that the presence of other interest as a motive increases how morally good actions are evaluated. Given that sympathy is associated with a motivation to help others it may be the case that sympathy simply increases how morally good an action is seen to be by providing the observer with information about a person’s desire to address the welfare or interests of others.

Thus, Studies 5a and 5b directly compare perceptions of actors who perform good behaviours out of sympathy, to actors who perform those behaviours out of a motivation to help. Higher goodness ratings of actions motivated by sympathy will be further evidence that sympathy increases how morally good an action is judged to be, beyond just informing the observer than an agent has a desire to help. In Study 5a sympathy is compared with a desire to help others using *volunteered* vignettes, and in
Study 5b sympathy is compared with a desire to help others using questioned vignettes\(^ {24} \).

**Method**

**Participants**

One hundred and thirty students (90 females, 40 males (\(M_{age} = 19.89, SD = 3.97\)) studying first or second year psychology at the University of Otago gave informed consent before participating in this study in exchange for partial course credit.

**Procedure and Materials**

This study used the same design as Study 4a, except that in the sympathy condition agents were described as being motivated by sympathy and a really strong desire to help, and in the no sympathy condition agents were described as having a really strong desire to help (See Table 4.3, Appendix G). Character judgements were also measured.

| **Table 4.3. Example volunteered vignettes for sympathy and no sympathy conditions.** |
|---------------------------------|---------------------------------|
| **Sympathy**                                      | **No Sympathy**                                      |
| Jackie donated $10 to help the victims of a flood. When asked why she donated the money she replied, “I felt a lot of sympathy for the victims and I had a really strong desire to help”. | Jackie donated $10 to help the victims of a flood. When asked why she donated the money she replied, “I had a really strong desire to help”. |

**Results**

Mean goodness scores were calculated by averaging across vignettes (\(\alpha = .83\)). Character scores were calculated by averaging the answer to the two character questions (\(\alpha = .79\)). Character scores were then averaged across vignettes (\(\alpha = .88\)). A 2 (sympathy, no sympathy; between) x 2 (order) ANOVA was conducted using goodness scores as the dependent variable. There was no significant difference between

\(^{24}\) Studies 5a and 5b could have been run as one study. However the sympathy condition is already within subjects and including both vignette conditions results in participants being asked to perform essentially the same task four times (as was done in Study 3). A similar problem arises if participants are presented with multiple conditions in a single block. It is prudent to avoid allowing participants to easily compare one condition to the other, a weakness that my studies already have (to a lesser degree) as sympathy conditions are presented sequentially.
judgements of moral goodness for the sympathy ($M = 5.75, SD = 1.00$) and the no sympathy conditions ($M = 5.80, SD = 0.99$), $F(1,129) = 0.84, p = .36, \eta^2_p < .01$. The order main effect, $F(1,129) = 0.27, p = .60, \eta^2_p < .01$, and the interaction were also not significant, $F(1,129) = 1.14, p = .23, \eta^2_p = .01$.

A 2 (sympathy, no sympathy; between) x 2 (order) ANOVA was also conducted using character scores as the dependent variable. There was no significant difference between judgements of character for the sympathy ($M = 5.95, SD = 0.77$) and the no sympathy conditions ($M = 5.90, SD = 0.77$), $F(1,129) = 1.50, p = .22, \eta^2_p = .01$. The order main effect, $F(1,129) = 1.63, p = .20, \eta^2_p = .01$, and the interaction were also not significant, $F(1,129) = 0.08, p = .78, \eta^2_p < .01$.

To test whether character judgements mediated the relationship between sympathy and judgements of moral goodness a bootstrap mediation analysis (Preacher & Hayes, 2004) was conducted (10000 bootstrap samples) to estimate the size of character mediation (paths $a \cdot b$; see Figure 3.2). As seen in Table 4.4, the analysis indicates that character judgements did not significantly mediate the relationship between sympathy and judgements of moral goodness, nor was the direct effect of sympathy (controlling for character; path $c'$) significant. However, while motive did not significantly predict character judgements, character judgements still significantly predicted goodness judgements.

<table>
<thead>
<tr>
<th>Table 4.4. Mediation analysis between sympathy condition and goodness judgements by character judgements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sympathy as a predictor</td>
</tr>
<tr>
<td>$B$</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Sympathy as a predictor</td>
</tr>
<tr>
<td>Note. * $p &lt; .05$. For model see Figure 3.2.</td>
</tr>
</tbody>
</table>

Discussion

Inconsistent with the findings of Study 4a, but consistent with Study 3 there was no significant difference between sympathy groups for the volunteered vignettes. Power was adequate to detect the influence of sympathy if we estimate power based upon
effect size in study 4a ($\eta^2_p = .07$, power$^{25} = .92$).

If we accept the findings of Study 4a that found significant differences between sympathy conditions for volunteered vignettes, then the failure to identify a difference in this study could be due to the change in comparison condition from “the right thing to do”, to “a really strong desire to help”, and indicates that sympathy and other interest could have an equivalent influence on goodness judgements.

However, another interpretation is that once the right kind of reason is present for performing a good action, the addition of subsequent right reasons does not influence goodness judgements of actions any further. This could explain why no effect of sympathy was identified in this study. If both sympathy and a strong desire to help are both the right kinds of reasons for performing morally good actions, then describing an action as being motivated by both will not influence goodness judgements any differently than if an action was described as being motivated by only one of them, as was the case in the no sympathy condition. And so the design rests on the assumption that influence of reasons on goodness judgements of actions are additive.

Study 5b
Similarly to Study 5a which looked at volunteered vignettes, and altered the sympathy and no sympathy conditions to include a desire to help, Study 5b looks at questioned vignettes, and describes agents in the sympathy condition as being motivated by sympathy and a really strong desire to help, and in the no sympathy condition, as having a really strong desire to help (See Table 4.5, Appendix H). Character judgements were also measured.

Method
Participants
One hundred and ten students (78 females, 32 males ($\text{Mage} = 23.41, \text{SD} = 5.81$) studying first or second year psychology at the University of Otago gave informed consent before participating in this study in exchange for partial course credit.

25 Correlation of between subjects factors = .70.
Procedure and Materials

The design was similar to Study 4b. The difference is that, the agent in the sympathy condition is motivated by sympathy and a strong desire to help, and in the no-sympathy condition is motivated by a strong desire to help (See Table 4.9, Appendix H).

Table 4.5. Example questioned vignettes for sympathy and no sympathy conditions

<table>
<thead>
<tr>
<th>Sympathy</th>
<th>No Sympathy</th>
</tr>
</thead>
</table>
| Matthew has a friend who was in a car accident and has been recovering in hospital for the last few weeks. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “I feel enormous sympathy for him and I have a really strong desire to help”. | Matthew has a friend who was in a car accident and has been recovering in hospital for the last few weeks. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “Not really, I just have a really strong desire to help”.

Results

Mean goodness scores were calculated by averaging across vignettes ($\alpha = .73$). Character scores were calculated by averaging the answer to the two character questions ($\alpha = .83$). Character scores were then averaged across vignettes ($\alpha = .84$). A 2 (sympathy, no sympathy; between) x 2 (order) ANOVA was conducted using goodness scores as the dependent variable. Goodness scores for the sympathy group were significantly higher ($M = 6.06, SD = 0.79$) than the no sympathy group ($M = 5.57, SD = 1.14$), $F(1,109) = 23.74, p = .000, \eta_p^2 = .18$. Neither the order main effect, $F(1,109) = 3.60, p = .06, \eta_p^2 = .03$, nor the interaction were significant, $F(1,109) = .66, p = .42, \eta_p^2 < .01$.

A 2 (sympathy, no sympathy; between) x 2 (order) ANOVA was conducted using character scores as the dependent variable. Character scores for the sympathy condition were significantly higher ($M = 6.02, SD = 0.74$) than the no sympathy group ($M = 5.60, SD = 0.94$), $F(1,109) = 32.13, p < .01, \eta_p^2 = .23$. The order main effect, $F(1,109) =1.27, p = .26, \eta_p^2 =.01$ was not significant, however the interaction between order and sympathy conditions was significant, $F(1,109) = 3.99, p < .05, \eta_p^2 = .04$. As seen in Table 4.6, the analysis indicates that character judgements significantly mediate
the relationship between sympathy and judgements of moral goodness, but that the direct effect of sympathy (controlling for character; path $c'$) remains significant.

Table 4.6. *Mediation analysis between sympathy condition and goodness judgements by character judgements*

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$c'$</th>
<th>$ab$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>B</td>
<td>SE</td>
<td>B</td>
</tr>
<tr>
<td>Sympathy as a predictor</td>
<td>.42*</td>
<td>.11</td>
<td>.62*</td>
<td>.07</td>
<td>.49*</td>
</tr>
</tbody>
</table>

* p < .05

*Discussion*

Unlike Study 5a, in which agents volunteered the reasons for their actions, Study 5b revealed differences between sympathy and no-sympathy conditions, as in Studies, 3, 4a and 4b. As with Study 4b mediation analysis found that character judgements partially mediated the relationship, but that sympathy remained a significant predictor when controlling for character. The results suggest, then, that although part of the difference between sympathy conditions is due to character judgements of agents, they also indicate that sympathy is a distinct reason from a desire to act in the interest of others.

Again a significant effect was found for the sympathy condition using *questioned* but not *volunteered* vignettes. *Volunteered* vignettes have proven more difficult to demonstrate an influence of sympathy on goodness judgements than *questioned* vignettes. The effect size for *questioned* vignettes in this study ($\eta_p^2 = .18$) was less than half the size observed in Study 4b ($\eta_p^2 = .38$). A reduction in effect size may be the reason why a significant effect was not observed for the *volunteered* vignettes in the previous study (Study 5a). If the effect size halved for *questioned* vignettes from Study 4b to this study then a crude estimate of the power of the previous study looking at *volunteered* vignettes by halving the power of Study 4a, gives a power estimate of ~ .65 for the previous study (Study 5a), which is not impressive and may be why an effect of sympathy condition was observed in this study for *questioned* vignettes, but not in the previous study for *volunteered* vignettes.
General discussion

The five studies in this chapter provide qualified support for the hypothesis that the presence of sympathy increases how morally good actions are evaluated to be. Sympathy was manipulated in two ways, agents were either asked if they acted out of sympathy, or agents volunteered sympathy as a reason for their actions. When asked if feeling sympathy was the reason why they performed a good action, the actions of agents who answered in the affirmative were rated more morally good than if the agent denied acting out of sympathy. As in Study 2 character judgements mediated the relationship between reasons and goodness judgements.

The evidence was less compelling that sympathy increases goodness judgements of actions when an agent reported without prompting that they acted out of sympathy. The influence of sympathy on goodness judgements was not observed for volunteered vignettes in Study 3. Study 4a found that when the outcomes of the actions performed by agents in Study 3 were less morally good, that sympathy influenced goodness judgements of actions, suggesting that a ceiling effect may have masked the influence of sympathy in Study 3. However no difference between sympathy conditions was again observed in Study 5a, when the reason for acting in the no sympathy condition was changed to a desire to help others.

That goodness judgements for questioned vignettes were more often influenced by sympathy, could be due to assumptions participants make about an agent’s reasons for acting in the no sympathy condition for volunteered vignettes. The challenge with these studies was to control the assumptions participants made about agents’ reasons for their actions. Inferring motives and reasons are fundamental to behaviour explanations (Malle, 2004), and there is evidence that this process is automatic (Hassin, Aarts, & Ferguson, 2005; Overwalle et al., 2012). The concern with volunteered vignettes is that in the no-sympathy condition participants could infer that the agent acted from sympathy because their actions helped others (a typical response to seeing others in need is sympathy).

Questioned vignettes were included to control for and test these assumptions, and that sympathy influenced goodness judgements for questioned vignettes and not volunteered vignettes is consistent with the concern that participants would infer sympathy from the actions of the agents in the no sympathy condition for volunteered
vignettes. However since sympathy was found to influence goodness judgements for volunteered vignettes in Study 4a, this account is less compelling, although this conclusion must be qualified as the contents of the vignettes in the volunteered and questioned conditions differed. In study 4a and 4b the sympathy manipulation influenced both volunteered and questioned vignettes, and it was found that the effect size for the influence of sympathy on volunteered vignettes was about half the size as for questioned vignettes. The overall results are more consistent with an interpretation that sympathy did not influence goodness judgements in Studies 3 and 5a because the effect size of the sympathy manipulation was relatively small, and the studies had low power to detect an effect of this magnitude.

However there is still the question of why the effect size was larger for questioned vignettes than volunteered vignettes. It may be the case that being questioned about sympathy has a larger influence on goodness judgements than volunteering sympathy because for questioned vignettes both the presence of sympathy increases, and the denial of sympathy decreases goodness judgements in the no sympathy condition. However we are not able to tell precisely what is going on here as one of the downsides of the questioned vignettes is the difficulty in identifying whether affirming sympathy or denying sympathy or both, influence goodness judgements. Future studies could examine whether denying sympathy (and the right kinds of reasons more generally) influence goodness judgements. It could also be the case that denying sympathy reduces goodness judgements because denying sympathy indicates that a person is not feeling the right kind of way when they are performing a good action. Studies 1-5 have examined the importance of being motivated by the right reasons when performing good actions, but it may also be important that a person feels the right way when performing good actions.

In Study 2 motives were found to influence goodness judgements, and these motives had a larger effect on goodness judgements (Study 2 self interest $d = 1.33$ self interest, $d = 1.26$ principle) than in the questioned sympathy conditions (Study 3 and 4b $d = 0.81$ and 0.82). This could be due to the types of motives used, but the difference in effect size could also be due to other differences between the study designs. It is probably more difficult to demonstrate an influence of reasons and motives on goodness judgements with the study design used in Studies 3-5, than the design of Study 2. Study
2 experimentally examined the influence of several motives on goodness judgements of actions. In this study the outcome of an agent’s actions was intended to be morally neutral, with only the presence or absence of motives influencing the moral goodness of the action. In Studies 3-5 the outcomes of the agent’s actions were morally good without further information about the agent’s motives, which may result in more difficulty identifying the influence of reasons on goodness judgements of actions because the influence of an agent’s reasons on goodness judgements is now one of several influences.

Conclusion
The studies in this chapter provide support for the hypothesis that sympathy is the right kind of reason for performing a morally good action, and that the presence of sympathy as a reason increases goodness judgements of actions. This evidence is consistent with findings from Study 2 that motives and reasons influence goodness judgements of actions. Also as in Study 2 it was found that the relationship between reasons and goodness judgements of actions was partially mediated by character judgements.

Collectively, studies 1-5 present evidence for the influence of specific motives and reasons on goodness judgements of actions. In the next chapter these findings, and specifically the finding that sympathy as a reason increases goodness judgements, are used to reinterpret the asymmetry effect (Pizarro, Uhlman, & Salovey, 2003) which has important consequences for the model of how people make goodness judgements of actions.
Chapter 5
Reasons, the asymmetry effect and simpler explanations

In this chapter I consider an important finding for a candidate model of goodness judgements. The asymmetry effect is an observation that blame but not praise is reduced for actions where the agent is deemed to have a loss of control over their actions. I hypothesize that differences in reasons for acting can account for the asymmetry effect, rather than differences in the morality of the actions. However, while exploring the alternative hypothesis a more parsimonious explanation for the asymmetry effect was identified and followed up in a series of studies.

One of the few findings in the literature specific to judgements of morally good actions is the asymmetry effect (Pizarro, Uhlman, & Salovey, 2003). Evidence suggests that when someone acts with reduced personal control (Alicke, 2000), the blame they receive for performing that action is reduced compared to if they performed the action deliberately. However, according to the asymmetry hypothesis, loss of control does not reduce praise for morally good actions (Pizarro et al., 2003). It has been argued that control information is not used in judgements of morally good actions because of assumptions observers make about an agent’s meta desires (Pizarro et al., 2003).

The asymmetry effect has important consequences for the development of a model of goodness judgements. Although no such model currently exists, studies – including those in this thesis – have identified aspects of an agent’s mental states, actions, and their outcomes that are likely to play a role. If one were to create a model of goodness judgements, blame models are an obvious place to start (Alicke, 2000; Guglielmo et al., 2009) as they build upon more general work on behaviour explanation.

The current models of blame minimally require a negative event to occur, and an agent to have a causal connection to the outcome (Alicke, 2000; Guglielmo, Monroe, & Malle, 2009; Shaver 1985 cited in Lagnado & Channon, 2008). Blame is then amplified or diminished by the intentionality of a person’s actions, a person’s motives and reasons, and the control a person has over their thoughts and actions (Alicke, 2000;
Guglielmo, Monroe, & Malle, 2009; Lagnado & Channon, 2008). The asymmetry effect has important consequences for the relevance of this model for goodness judgements, as it casts doubt on the relevance of an agent’s control over their actions while performing a morally good action.

**The Asymmetry effect**
Pizarro, Uhlmann and Salovey (2003) examined the influence of loss of control on judgements of praise and blame. Previous research has identified that a person who is judged to have reduced control over their actions often receives less blame (Alicke, 2000). For example, crimes of passion are often judged less harshly than crimes committed in a deliberate and rational manner (Pizarro et al., 2003).

Loss of control comes in many forms. According to Alicke’s (2000) culpable control model, people evaluate control at several stages: volitional behaviour control, which involves assessing intentionality or the degree to which a person wanted to act as they did; causal control, the degree to which the agent’s behaviour caused the outcome; and volitional outcome control, the degree to which the outcome was desired and foreseen. Pizarro et al. (2003) investigated loss of control from an internal source; for example, mental illness or extreme emotional episodes are examples of internal forces that can compel a person to act with reduced control (which Alicke would call volitional behaviour control). This kind of loss of control is also referred to as diminished capacity (Guglielmo, Monroe, & Malle, 2009).

Pizarro et al. (2003) argue that while a loss of control leads to reduced blame for morally bad actions this is not the case for praise and morally good actions. To test this hypothesis over three studies Pizarro et al. presented participants with vignettes, manipulating whether the action of the agent in each vignette was morally good or bad, and whether the agent performed the action deliberately and intentionally or because of an overwhelming and uncontrollable emotion (Table 5.1).

---

26 These models look at blame and not judgements of moral transgressions. Though these are similar concepts there are some important differences. Cushman (2008) found that blame was sensitive to mental states and outcomes, as specified in these models, but wrongness judgements are more reliant on mental states (or motives). Analogously we may find that praise for actions is more closely associated with the outcome of the actions than goodness judgements.
Table 5.1. *Sample vignettes from Pizarro, Uhlmann & Salovey (2003)*

<table>
<thead>
<tr>
<th>Valence</th>
<th>Deliberate</th>
<th>Impulsive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Jack deliberately and intentionally gave the homeless man his only jacket, even though it was freezing outside.</td>
<td>Because of his overwhelming and uncontrollable sympathy, Jack impulsively gave the homeless man his only jacket even though it was freezing outside.</td>
</tr>
<tr>
<td>Negative</td>
<td>Jack calmly and deliberately smashed the window of the car parked in front of him, because it was parked too close to his.</td>
<td>Because of his overwhelming and uncontrollable anger, Jack impulsively smashed the window of the car parked in front of him because it was parked too close to his.</td>
</tr>
</tbody>
</table>

Supporting their hypothesis, the authors found that participants evaluated the agents who deliberately performed morally bad actions as more blameworthy than agents who performed the same actions because of uncontrollable anger. Furthermore, they found that participants judged agents who deliberately performed morally good actions just as praiseworthy as those who performed morally good actions because of an uncontrollable sympathy.

Pizarro et al. went on to hypothesise that the asymmetry is due to assumptions observers make about people’s second-order desires (or “meta-desires”). According to the authors, second-order desires are an individual’s acceptance or rejection of a desire or impulse. For example, for a drug addict “there is a second-order desire (not to be an addict) and a first-order desire (a compulsion to continue to take drugs)” (Pizarro et al., p.267). The authors argue that if an agent’s second-order desires are consistent with their actions, then loss of control will not diminish the amount of praise or blame an agent receives for their actions.
An alternative explanation

Throughout this thesis I have presented evidence for the influence of motives and reasons on goodness judgements of actions. In Studies 3-5, I argued that sympathy as a reason for an agent’s actions increases how morally good an observer judges the agent’s actions to be. Based on these findings I propose an alternative explanation for the asymmetry effect, in terms of the influence of reasons on goodness judgements of actions.

As noted, evidence for the asymmetry comes from several studies in which participants were presented with vignettes describing agents performing either morally good or bad actions (Pizarro et al., 2003). The agents acted either deliberately or because of an uncontrollable emotion. When an agent uncontrollably performs a morally bad action they do so because of uncontrollable anger, and when they perform a morally good action they do so because of uncontrollable sympathy. It was found that blame is discounted for actions resulting from uncontrollable anger, but that praise is not discounted for actions resulting from uncontrollable sympathy.

I propose that praise is not discounted for actions performed because of uncontrollable sympathy because sympathy is the right kind of reason for performing a good action. As I showed in Studies 3-5, the presence of sympathy is sufficient to increase the goodness of an action. It may, therefore, be that praise is not discounted in the uncontrollable sympathy condition because the presence of sympathy increases the moral goodness of the action, masking the fact that the uncontrollable nature of the action would otherwise have reduced praise. However, a closer look at the evidence reveals that the asymmetry effect can be explained in an even simpler way.

A more parsimonious explanation

In Pizarro et al. (2003) when the agent uncontrollably performs a morally bad action they do so because of uncontrollable anger and when an agent performs a morally good action they do so because of uncontrollable sympathy. Rather than the asymmetry in reduction of praise and blame being due to the morality of the agent’s actions, it may be that differences in the uncontrollable emotions account for the asymmetry. In Pizarro et al. uncontrollable anger and sympathy are used to reduce the perceived control an agent

27 "A first principle not formally recognized by scientific methodologists: when you run into something interesting, drop everything else and study it." (p. 223, Skinner, 1956)
has over their actions. If it is the case that the reduction of control differs between these overwhelming emotions then the asymmetry effect may be due to differences in control rather than the morality of the agent’s actions.

There are reasons to suspect that anger and sympathy differ in the extent to which they are perceived to reduce control. A person losing their temper and control of themselves is an ordinary sort of explanation for behaviour. Rage is a common explanation for why people act in ways they would not act with a cooler head (Lakoff & Koveceses, 1987; Lupton, 2002). However, although people can experience a great deal of sympathy, experiencing sympathy to a degree that a person is perceived to have reduced control over their actions is not commonly associated with sympathy. If people have much more experience with anger leading to a loss of control they may be more willing to accept that a person has reduced control over their actions if the loss of control was due to anger rather than sympathy. Participants may not believe that overwhelming sympathy reduces control to the same degree as overwhelming anger, and describing an action as the result of uncontrollable sympathy does not guarantee that people will treat the action as uncontrolled. In the following studies I examine whether the asymmetry effect is the result of a confound between emotion and morality.

Is the asymmetry due to different effects of anger and sympathy on perception of control?

If uncontrollable anger is perceived to reduce control to a greater degree than uncontrollable sympathy, then the asymmetry effect should disappear if the same uncontrollable impulse is responsible for an agent performing both morally good and bad actions. In Study 6 I attempt to equate control by having the same kind of impulse responsible for an agent performing both morally good and morally bad actions.

The study uses a design similar to that used by Pizarro et al. (2003), presenting vignettes in which agents perform morally good or bad actions either deliberately or impulsively. In order to keep the loss of control consistent between conditions, instead of the impulse resulting from overwhelming anger or sympathy, the impulse is the result of frontal lobe damage or a history of impulse control problems, which both fall under Pizarro et al.’s category of “internal impulse”. I hypothesised that when loss of control

---

28 As evidence for this statement, besides Pizarro et al., (2003) I could not find another study discussing sympathy and loss of control, while there are numerous studies and articles discussing anger and loss of control.
due to an internal impulse is consistent between morally good and bad actions the asymmetry will disappear. Specifically, I predicted that when the impulse is caused by either frontal lobe damage or a history of impulse control difficulties, impulsive morally bad actions will be judged to be less bad than deliberate morally bad actions, and impulsive morally good actions will be judged less good than deliberate morally good actions.

The asymmetry hypothesis applies to both judgements or praise/blame and moral goodness/badness, right and wrong (Pizarro et al., 2003). There is evidence for differences between judgements of blame and wrongdoing. For example, Cushman (2008) found that judgements of wrongness are more sensitive to information about intent, whereas blame judgements are more sensitive to information about outcome. Nevertheless, I will discuss the following research in terms of judgements of morally good and bad actions to be consistent with the rest of this thesis.

Study 6

Method

Participants

One hundred and fourteen (92 female, 22 male; $\text{Mage} = 21.82$, $SD = 7.64$) participants studying first or second year psychology at The University of Otago gave informed consent before participating in this study in exchange for partial course credit.

Procedure and Materials

The questions relevant to this study were embedded in a larger study looking at participants’ moral judgements and assembled, along with an additional, unrelated study into a single paper and pencil questionnaire. The questionnaire was administered to groups of approximately 3 to 20 participants. Participants gave informed consent before they began the study.

Participants were given written instructions (Appendix I) informing them of the researcher’s interest in moral intuitions, and that they would read a description of someone performing an action and then be asked three questions. Participants were then instructed to circle the number on the scale to indicate their answer to each question.

Two pairs of vignettes from the questionnaire are relevant to this study. Both
pairs manipulate control by describing a person performing morally good or bad actions, having acted either deliberately or because of a powerful impulse. The vignettes describing morally good actions were inspired by real life, self reported accounts of good deeds (http://www.thegooddeedsorganisation.com) and the morally bad actions were taken from examples in other studies as well as news events. In the first manipulation, an impulse is the result of frontal lobe damage. In the second manipulation the actor has a history of impulse control problems.

*Manipulation 1: Frontal lobe damage*

For the frontal lobe manipulation participants read and responded to two vignettes (Table 5.2 and Appendix I). In the first vignette the actor performed a morally bad action and in the second vignette the actor performed a morally good action. Whether the actor was described as having performed the action deliberately or due to an impulse was manipulated between subjects. When an action was described as having been performed deliberately it was in the form: *X deliberately and intentionally did Y*. An action performed impulsively took the form: *X has frontal lobe damage, which results in uncontrollable impulsive behaviour. Because of an overwhelming and uncontrollable impulse, X impulsively [did] Y*. After reading each vignette participants answered three questions, each accompanied by a nine-point scale. Two moral judgement questions: *How moral was it that X [did] Y?* [-4(Extremely morally bad) to 4(Extremely morally good)], *How much blame or praise should X get for Y?* [-4(Extreme blame) to 4(Extreme praise)], and one control question as a manipulation check, *How much control do you think X had over his/her actions when he/she [did] Y?* [0(No control) to 8(Complete control)].
Table 5.2. Example vignettes for frontal lobe damage manipulation

Control manipulation (between)

<table>
<thead>
<tr>
<th>Morality (within)</th>
<th>Deliberate</th>
<th>Impulsive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morally Good</td>
<td>Jeffrey deliberately and intentionally gave half his savings ($2000) to charity.</td>
<td>Jeffrey has frontal lobe damage, which results in uncontrollable impulsive behaviour. Because of an overwhelming and uncontrollable impulse, Jeffrey impulsively gave half his savings ($2000) to charity.</td>
</tr>
<tr>
<td>Morally Bad</td>
<td>Charlotte deliberately and intentionally punched a student in her class who was rude to her.</td>
<td>Charlotte has frontal lobe damage which results in occasional overwhelming and uncontrollable impulses. Because of one such overwhelming and uncontrollable impulse Charlotte punches a student in her class who is rude to her.</td>
</tr>
</tbody>
</table>

Manipulation 2: Impulse history

The impulse history manipulation was similar to the frontal lobe manipulation, except that instead of an impulse caused by frontal lobe damage, the agent in the impulse condition was described as having a history of impulse control problems. Vignettes for the impulse condition took the form: *X has had impulse control difficulties since he/she was a child. He/she occasionally experiences extremely powerful and compelling urges to do things that he/she can’t control. Because of one such overwhelming and uncontrollable impulse, X impulsively [did] Y* (Table 5.3, Appendix I). Note that as with the frontal lobe manipulation, participants read one vignette with a morally good outcome and one vignette with a morally bad outcome. However this manipulation differed as rather than both good and bad vignettes being either impulsive or deliberate, participants saw one impulsive and one deliberate vignette.
Table 5.3. Example vignettes for impulse control difficulty manipulation

<table>
<thead>
<tr>
<th>Morality (within)</th>
<th>Control manipulation (between)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deliberate</td>
</tr>
<tr>
<td>Morally Good</td>
<td>While waiting in line at a checkout to pay for her groceries Hilda notices that the lady in front of her doesn’t have enough money to pay for her groceries. Hilda deliberately and intentionally gives the lady $50 to pay for her groceries.</td>
</tr>
<tr>
<td>Morally Bad</td>
<td>Jenna deliberately and intentionally left the restaurant without paying the bill.</td>
</tr>
</tbody>
</table>

Results

Manipulation 1: Frontal lobe damage

Following Pizarro et al., moral judgement scores were calculated by taking the absolute values for participants’ responses to the two moral judgement questions. This allowed for a comparison between the extremity of values for both morally good and bad actions. The two moral judgement questions were then averaged for each vignette ($\alpha = .94$). The data from four participants was discarded because their responses indicated that they did not consider the actions of the agent in the morally good vignette to be

---

29 For a replication of the asymmetry effect see Appendix J.
A 2(morality: morally good, bad; within subject) x 2(control manipulation: deliberate, impulsive; between subject) ANOVA was run. The main effect for the control manipulation was significant, $F(1,113) = 141.16, p < 0.01, \eta^2_p = .56$, while the main effect for morality was not, $F(1,113) = 2.43, p = .12, \eta^2_p = .02$. The interaction between morality and control manipulation was also significant, $F(1,113) = 7.87, p < .01, \eta^2_p = .07$. Follow up $t$ tests ($\alpha = 0.05$) revealed that the mean moral judgement score was significantly larger for deliberate morally bad actions ($M = 3.56, SD = 0.63$) than for impulsive morally bad actions ($M = 1.72, SD = 0.88$), $t(113) = 12.90, p < .01, 95\% CI [1.56, 2.12], d = 2.40$. It was also the case that the mean moral judgement score for deliberate morally good actions ($M = 3.45, SD = 0.66$) was significantly larger than the mean for impulsive morally good actions ($M = 2.10, SD = 1.20$), $t(116) = 7.57, 95\% CI [1.00, 1.71], d = 1.38$ (Figure 5.1).

Figure 5.1. Frontal lobe manipulation. Mean moral judgement score ($\pm SE$) for morally good and bad acts that were either deliberate or due to an impulse.
12.80, \( p < .01, \eta_p^2 = .01 \) as was the control manipulation, \( F(1,112) = 599.41, p = .01, \eta_p^2 = .84 \). The interaction between morality condition and control manipulation was not significant \( F(1,112) = 1.21, p = .27, \eta_p^2 = .01 \). Control scores for deliberate morally good actions (\( M = 6.95, SD = 2.03 \)) were judged to be performed with significantly more control than actions that resulted from an impulse (\( M = 1.05, SD = 1.33 \)), \( t(115) = 18.71, p < .01, d = 3.44 \). Similarly control scores for deliberate morally bad actions (\( M = 7.61, SD = 0.65 \)) were judged to be performed with significantly more control than actions that resulted from an impulse (\( M = 1.44, SD = 1.69 \)), \( t(116) = 25.87, p < .01, d = 4.82 \).

**Manipulation 2: Impulse control history**

Moral judgement scores were calculated in the same way as for the frontal lobe manipulation (\( \alpha = .81 \)). Unlike the frontal lobe manipulation participants did not see both deliberate or both impulsive vignettes, they saw one deliberate and one impulsive vignette (Table 5.4), so a 2 x 2 ANOVA could not be conducted.

**Table 5.4. Questionnaire versions**

<table>
<thead>
<tr>
<th>Outcome (within)</th>
<th>Version 1</th>
<th>Version 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morally Bad</td>
<td>Deliberate</td>
<td>Impulsive</td>
</tr>
<tr>
<td>Morally Good</td>
<td>Impulsive</td>
<td>Deliberate</td>
</tr>
</tbody>
</table>

*Note.* Participants received either questionnaire version 1 or 2.

However, an analysis of simple effects revealed that deliberate morally bad actions (\( M = 3.31, SD = 0.82 \)) had significantly greater mean moral judgement scores than impulsive morally bad actions (\( M = 1.81, SD = 1.03 \)), \( t(118) = 8.95, p < .01, 95\% CI [1.18, 1.84], d = 1.60 \). It was also the case that deliberate morally good actions (\( M = 3.08, SD = 0.89 \)) received significantly higher mean moral judgement scores than impulsive morally good actions (\( M = 2.55, SD = 1.03 \)), \( t(119) = 3.02, p < .01, 95\% CI [1.18, 1.84], d = 1.60 \).
CI[0.18, 0.88], $d = 0.55$, (Figure 5.2).

**Figure 5.2.** Impulse control history manipulation. Mean moral judgement score ($\pm$SE) for morally good and bad acts that were either deliberate or due to an impulse.

Control scores for deliberate morally good actions ($M = 7.33$, $SD = 1.03$) were judged to be performed with significantly more control than actions that resulted from an impulse ($M = 2.79$, $SD = 2.14$), $t(119) = 14.63$, $p < .01$, $d = 2.70$. Similarly control scores for deliberate morally bad actions ($M = 7.71$, $SD = 0.68$) were judged to be performed with significantly more control than actions that resulted from an impulse ($M = 1.93$, $SD = 1.78$), $t(116) = 23.92$, $p < .01$, $d = 4.29$.

**Discussion**
Consistent with the asymmetry hypothesis, deliberate morally bad actions were judged to be morally worse than actions that resulted from an uncontrollable impulse. However, inconsistent with the asymmetry hypothesis it was found in both frontal lobe and impulse history manipulations, that deliberate morally good actions were judged to be more morally good than impulsive morally good actions. The manipulation checks indicated that control was judged to be significantly greater for the deliberate condition.
for both good and bad actions. The results are consistent with the hypothesis that when the reduction of control is consistent across morally good and bad actions, impulsive morally bad actions will be judged less bad than deliberate morally bad actions, and impulsive morally good actions will be judged less good than deliberate morally good actions.

In the frontal lobe manipulation there was a significant interaction between morality and control manipulation, such that lack of control moderated morality slightly more for negative than for positive actions. Although this difference in sensitivity is arguably consistent with a weak version of the asymmetry hypothesis, it is disconfirms the strong version (advocated by Pizzaro et al., 2003) in which judgements of moral goodness are not influenced by control information at all.

A limitation of the frontal lobe manipulation is that only one vignette was run for each morality condition and the order was not randomised. A reduction in how morally good impulsive actions were judged to be may be due to an order effect, or the specific vignette used. Furthermore, participants in the impulse condition all read a vignette depicting a morally wrong action first. If participants had just reduced how morally wrong they thought the agent’s actions were because the agent had acted impulsively, it may be that the idea of reducing the extremity of a moral judgement because of loss of control may be more accessible when judging the positive behavior, thereby inflating the difference between control conditions. It could also be that by placing these two similar vignettes in close proximity, participants may have been given the expectation that they should reduce how morally good they judge the impulsive action to be, relative to the deliberate action. This interpretation is weakened, however, by the results of the impulse control history manipulation, which used a different order, in which presentation of the impulsive morally good vignette was preceded by the deliberate morally bad vignette, and not the impulsive morally bad vignette.

A further limitation was the inclusion of the control question. Explicitly asking about control could draw the participant’s attention to control information, or suggest to participants that they should use the control information in the vignettes to make their judgements.

Study 7
Study 6 provides evidence that when causes of positive and negative behaviours are
equated in terms of their controllability, loss of control can reduce praise for morally good actions, contrary to the asymmetry hypothesis proposed by Pizarro et al. (2003). Study 6 however, was limited by the use of a small number of vignettes and inadequate randomisation. Study 7, therefore, was designed to replicate the findings of Study 6, with a larger number of randomised vignettes. Additionally, participants’ perceptions of each agent’s meta-desires will be measured to address the meta-desire interpretation of the asymmetry effect.

Method
Participants
Sixty eight (40 female, 28 male; Mage = 24.33, SD = 7.10) studying first or second year psychology at The University of Otago gave informed consent before participating in this study in exchange for partial course credit.

Materials and Procedure
This study was one among three unrelated studies completed by participants. Stimuli were presented using SuperLab 4 experiment presentation software on iMac 21-inch computers in private experimental cubicles. Participants read vignettes and responded using the keyboard. The dependent measures were embedded among several other questions, reported later.

Participants were given similar instructions to Study 6 (Appendix K). Depending on which control manipulation was presented first, participants read four vignettes describing actions that were all performed either deliberately or all performed impulsively (counterbalanced). As in the first manipulation in Study 6, impulsive actions were described as caused by impulses resulting from frontal lobe damage. The vignettes depicted two morally good, and two morally bad actions. Participants then completed an unrelated filler task, after which they were instructed that they would read about the same situations again, but this time the person in the vignettes had different reasons for their actions. Participants were then presented with vignettes from the other control condition. The design of the study was thus a 2 (morality: morally good, morally bad; within-subject) x 2 (control manipulation: deliberate, impulsive; within-subject) x 2 (vignette; within-subjects) x 2 (order of control conditions; between-subjects). After reading each vignette participants were asked three questions (modelled on those used
by Pizarro et al., 2003): How much praise or blame should X get for Y? How morally
good or bad was it that X [did] Y? How positively or negatively should X be evaluated
for Y? The responses scale differed from Study 6 as 13 point scale was used (-
6(Extremely morally bad) to 6(Extremely morally good)) instead of a 9 point scale.
Participants also answered two questions used in Pizarro et al. (2003, Study 3) to assess
meta-desires: To what extent did X really want to do Y [-3(really did not want to do it),
0 (neither), 3 (really wanted to do it)]; To what extent do you think X wanted to have an
impulse to Y [-3 (really did not want to have the impulse), - 0 (neither), – 3 (really
wanted to have the impulse)]

Results
Moral judgement scores were calculated as in Study 6. Nineteen data points (out of 544)
were discarded from participants who judged “bad” actions to be good, or vice versa.
Mean moral judgement scores were calculated for each vignette by averaging the
answers to the three questions asked after each vignette (α = .97). Mean moral
judgement scores were calculated by averaging the moral judgement scores across
vignette (α = .92) within each condition. A 2(morality: morally good, morally bad;
within-subject) x 2(control manipulation: deliberate, impulsive; within-subject) x
2(vignette; within-subjects) x 2(order; between-subjects) ANOVA was conducted.

The main effect for vignette was not significant F(1,53) = 2.86, p = .10, ηp² < .01. All other main effects were significant; control manipulation F(1,53) = 168.56, p < .01, ηp² = .76, morality F(1,53) = 21.84, p < .01, ηp² = .10, and order F(1,53) = 6.53, p = .01, ηp² = .11. Follow up t tests found that deliberate morally wrong actions (M = 4.33
SD = 1.21) were judged to be more morally wrong than impulsive morally bad actions
(M = 2.34, SD = 1.20), t(65) = 11.887, p < .01, 95% CI [1.66, 2.33], d = 1.65, and that
deliberate morally good actions (M = 4.90, SD = 1.15) were judged as significantly
more morally good than impulsive morally good actions (M = 3.33, SD = 1.48), t(66) =
9.1443, p < .01, 95% CI[1.22, 1.91], d = 1.18 (Figure 5.3).

There were significant two way interactions between vignette and morality,
F(1,53) = 25.382, p < 0.01, ηp² = 0.32, which indicates that the difference between

30 I did not include question 1 used by Pizarro et al. Study 3 (This person would rather not have an
impulse to Y) in order to reduce the number of questions participants answered. It was also a negatively
worded version of one of the questions, which could easily be omitted.
morality judgements for the two morally good vignettes differed from the difference between the two morally bad vignettes, and order and control manipulation, $F(1,53) = 6.33, p = .02, \eta_p^2 = 0.15$. This indicates that when the impulsive control manipulation was presented first the difference between impulsive and deliberate conditions was larger than when the deliberate condition was presented first. No other effects reached significance. The full results appear in Appendix L.

![Figure 5.3](image)

Figure 5.3. Study 7. Mean moral judgement scores (±SE) for morally good and bad acts that were either deliberate or due to an impulse.

The two questions assessing meta-desires were averaged ($\alpha = .92$). A 2 (morality: morally good, morally bad; within-subject) x 2 (control manipulation: deliberate, impulsive; within-subject) x 2 (vignette; within-subjects) x 2 (order; between-subjects) ANOVA was conducted. The vignette and order main effects were not significant. However the morality, $F(1,66) = 48.52, p < .01, \eta_p^2 = .42$, and control, $F(1,66) = 152.35, p < .01, \eta_p^2 = .70$, main effects were significant. Follow up $t$ tests found that meta-desire scores were significantly higher for deliberate morally bad actions ($M = 1.06, SD = 1.40$) compared to impulsive morally bad actions ($M = -0.79, SD = 1.31$), $t(67) = 8.57, p < .01, d = 1.46$. Meta-desire scores were also significantly
higher for deliberate morally good actions \((M = 2.14, SD = 0.89)\) when compared with impulsive morally good actions \((M = 0.08, SD = 1.19)\), \((67) = 13.60, p < .01, d = 1.96\), (Figure 5.4).

The only significant interactions were the three way interaction between morality, control and vignette, \(F(1,66) = 10.13, p < .01, \eta_p^2 = .13\), which indicated that though the overall morality x control interaction was not significant, this interaction differed between vignettes, and the three way interaction between morality, control manipulation and order, \(F(1,66) = 4.34, p = .04, \eta_p^2 = .06\), which indicated again, though the morality x control interaction was not significant, the order in which the control condition was presented influenced this interaction. The full results appear in Appendix L.

![Figure 5.4. Meta-desire scores for morally good and bad, deliberate and impulsive conditions.](image)

**Discussion**

The results replicate the finding from Study 6 that moral judgement scores for both impulsive morally good and bad actions were significantly lower than for deliberate morally good and bad actions. Inconsistent with the frontal lobe manipulation in Study
there was no significant interaction between control manipulation and morality, suggesting similar sensitivity to control information for both morally good and bad actions. The power to detect the interaction of the size identified in the Study 6 frontal lobe manipulation ($\eta_p^2 = .07$) was .97. The fact that an interaction was not identified in this study, suggests that if there is a sensitivity difference between the use of control information for judgements of good and bad actions, the effect size of the difference is quite small.

It was also found that participants thought deliberate actions were more consistent with an agent’s meta-desires than impulsive actions, and that good action were more consistent than bad actions with the agent’s meta-desires. Further examination of meta-desires and these results will be made in the general discussion.

**Summary of Studies 6 and 7**

Studies 6 and 7 provide evidence that people are sensitive to control information for both morally good and bad actions, and that reduced control leads to a reduction in the extremity of moral judgements for both morally good and bad actions. These findings suggest that praise was similar for deliberate and impulsive actions in Pizarro et al. (2003) because participants did not think uncontrollable sympathy reduced an agent’s control.

It is possible, however, that the results of Studies 6 and 7 are inconsistent with the asymmetry hypothesis because the kinds of impulses in Studies 6 and 7 differed from those used in Pizarro et al. (2003). Guglielmo et al. (2009) found that blame mitigation occurs most strongly when control (choice, in the author’s words) was reduced by factors such as brain tumours or schizophrenia, and that mitigation occurred to a lesser degree when overwhelming emotion was responsible for the loss of control (Guglielmo et al., 2009). So perhaps the asymmetry hypothesis could be restated more narrowly as: control information is not used to discount goodness judgements when the loss of control is sufficiently small.

**Convergent evidence**

Ideally, unconfounding Pizarro et al.’s (2003) study would involve fully crossing anger, sympathy and morality. This would mean running a study that examined both morally
good and bad actions that are the result of uncontrollable anger as well uncontrollable
sympathy. The fully crossed design would be 3 (control manipulation: anger, sympathy,
deliberate) × 2 (morality: morally good, morally bad). Ideally similar vignettes would
be used for each combination of the two factors.

This design was run; however, the control manipulation didn’ t influence any of
the conditions and did not replicate the basic asymmetry effect for anger and sympathy.
This may have been due to the difficulty in constructing vignettes for this design.
Constructing vignettes where a person performs a morally bad action because of
sympathy without making the vignette too different from the comparison vignette where
the agent acts out of anger was problematic. Sympathy implies concern for someone or
something, and the situations I devised, where the agent performed a morally bad action
because of overwhelming sympathy, were infused with noble sympathetic intentions.
This made interpreting results only considering control differences difficult. To aid
plausibility, the vignettes turned out to be somewhat ambiguous as to the agent’s
motivation.

Given the difficulty implementing a fully crossed design the next step was to
generate data on specific combinations of the factors that would speak to the falsity of
the asymmetry hypothesis. The original Pizarro et al. study has 4 of the 6 cells of a
completely crossed design (Table 5.5), and so it confounds anger with morally bad
actions and sympathy with morally good actions. Evidence for falsification of the
asymmetry hypothesis can be found by considering data from either of the missing
cells. Anger × morally good action and sympathy × morally bad action.

Table 5.5. Full design to test the asymmetry hypothesis

<table>
<thead>
<tr>
<th>Morality</th>
<th>Deliberate</th>
<th>Anger</th>
<th>Sympathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morally Good</td>
<td>Praise</td>
<td>? (study 8)</td>
<td>Praise</td>
</tr>
<tr>
<td>Morally Bad</td>
<td>Blame</td>
<td>Discounted blame</td>
<td>?</td>
</tr>
</tbody>
</table>

*Note. Combinations not run by Pizarro et al (2003) indicated by ‘?’.* The asymmetry hypothesis predicts praise for anger × morally good, and discounted blame for sympathy × morally bad. Study 8 investigates anger × morally good.
The following study considers evidence for falsification of the asymmetry hypothesis by obtaining data on morally good actions resulting from uncontrollable anger. The asymmetry hypothesis predicts that judgements of moral goodness will be identical for deliberate morally good actions and impulsive morally good actions resulting from uncontrollable anger. If uncontrollable anger reduces judgements of moral goodness then there is evidence that control information is used for judgements of morally good actions, that sympathy does not reduce control in the same way as anger, and that the asymmetry hypothesis is false.

Study 8

Method

Participants
One hundred and fourteen participants studying first or second year psychology at the University of Otago gave informed consent before participating in this study in exchange for partial course credit.

Materials and Procedure
This study was one among three unrelated studies completed by participants. Stimuli were presented using SuperLab 4 experiment presentation software on iMac 21-inch computers in private experimental cubicles. Participants read vignettes and responded using the keyboard. Participants read and evaluated a set of three vignettes twice. Each set of vignettes described a person performing morally good actions (Table 5.6 and Appendix M). In the anger condition the agent was described as having performed the action because of overwhelming uncontrollable anger. Vignettes were of the form: *Because of overwhelming and uncontrollable anger at Z, X impulsively did Y.* In the deliberate condition participants read the same vignettes but this time the action was described as being performed deliberately and intentionally. Vignettes took the form: *X deliberately and intentionally [did] Y.* After each vignette participants were asked three questions, as in Pizzaro et al. (2003): *How praiseworthy was it that X [did] Y? How morally good was it that X [did] Y? How positively should X be evaluated for Y?*

Participants were instructed that they would be presented with several situations and asked what they thought about the actions of the people in each situation. They
were then provided with an example question and an example of one of the scales they would be using (see Appendix M). Participants then completed either the *uncontrollable anger* or *deliberate* condition. Between each condition there was a break and participants were instructed that they would see similar but different vignettes again and that they should read them carefully. Participants were then presented with vignettes from the other control condition. The order in which participants completed the control conditions was counterbalanced.

**Table 5.6. Example stimuli for control manipulations.**

<table>
<thead>
<tr>
<th>Morality</th>
<th>Deliberate</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morally Good</td>
<td>Betty discovered that the company she worked for was dumping chemicals illegally into a local stream. Betty deliberately and intentionally reported the company to the police.</td>
<td>Because of her overwhelming and uncontrollable anger at discovering the company she worked for was dumping chemicals illegally into a local stream, Betty impulsively reported the company to the police.</td>
</tr>
</tbody>
</table>

**Results**

Mean moral judgement scores were calculated for each vignette by averaging the answers to the three questions asked after each vignette ($\alpha = .88$). Due to the low internal consistency between vignettes ($\alpha = .44$) they were included as an additional within-subjects factor in the following analysis.

A 2 (control manipulation: anger, deliberate; within-subjects) x 3 (vignette; within-subjects) x 2 (order; between-subjects) ANOVA was run. Neither vignette, $F(2,224) = .79, p = .46, \eta^2_p = .001$, nor order, $F(1,112) = 0.02, p = .90, \eta^2_p < .00$, main effects were significant. However the control manipulation main effect was significant, with the deliberate condition ($M = 5.94, SD = .74$) receiving significantly higher mean moral goodness ratings than the anger condition ($M = 5.35, SD = 1.10$), $F(1,112) = 58.00, p < .01, \eta^2_p = .345$. 
The three way interaction, control manipulation x vignette x order was significant, $F(2,224) = 5.47$, $p < .01$, $\eta^2_p = .05$, as was the two way integration between control and order, $F(1,112) = 4.66$, $p = .03$, $\eta^2_p = .04$, and seems to be driven by vignette 1 in order 1 (Figure 5.5). The two way interactions control x vignette, $F(2,224) = 2.49$, $p = .09$, $\eta^2_p = .02$, and vignette x order $F(2,224) = 1.04$, $p = .35$, $\eta^2_p = .01$ were not significant.

![Image](image.png)

**Figure 5.5.** Mean moral judgement scores for deliberate and anger conditions displayed by vignette and order.

**Discussion**

The control manipulation main effect indicates that participants judged deliberate morally good actions to be significantly more morally good than morally good actions that resulted from uncontrollable anger.

There was a significant three-way interaction between control manipulation, vignette, and order, which seems to be driven by one of the vignettes in the deliberate condition. This vignette does not seem to be driving the main effect though, as it is actually reducing the magnitude of the control manipulation main effect. There was also a significant interaction between control manipulation and order that seems to be driven by the same mean as the three-way interaction.
These results are inconsistent with the asymmetry hypothesis, and suggest that in the original Pizarro et al. study blame was reduced for morally bad actions, and praise was not for morally good actions, not because of differences in the morality of the actions, but because uncontrollable anger reduces control while uncontrollable sympathy does not.

Another interpretation could be that the anger expressed by the agents in each vignette results in the agents being judged negatively by the observer, and as a consequence the actions they performed are judged to be less morally good. In this interpretation, judgements of moral goodness are not reduced because uncontrollable anger reduces the control a person has over their actions, but because uncontrollable anger does not reflect well on the agent and they are seen as a worse person. Spontaneous evaluations of people influences blame judgements (Alicke, 2000), and Studies 2-5 have shown the influence of character judgements on evaluations of actions. This interpretation requires that the anger displayed by the agent results in them being evaluated negatively, even though being justifiably angry or expressing moral outrage could also reflect well on a person’s character.

Still another interpretation is that when performing a morally good action out of anger, the agent is seen to be doing the action to punish the person or thing with which they are angry. In the example stimuli in Table 5.6., the agent is angry at the company for dumping chemicals, and so she calls the police. Participants may interpret the situation as the agent being motivated to punish the company rather than to help others/the environment by reporting the company. Knobe (2003) has shown that praise is reduced if a morally good action is a side effect of the main reason the person acted (Knobe, 2003).

An argument could also be made that the agent in the anger condition did not have the right kind of motive or reason for performing a good action. I have argued in Studies 1-5 that goodness judgements of actions increase if an agent has the right kind of reason for performing the action. In this case, perhaps, agent’s actions in the anger condition were judged to be less good than in the deliberate condition because anger is not the right kind of reason for performing a good action.
Additional data

These alternatives are plausible; however, the argument that anger is particularly effective in reducing control, and consequently moral judgements, is strengthened by additional data from Studies 6 and 7. Participants in those studies were also asked questions about uncontrollable emotions other than anger, such as disgust and fear, which led to agents performing morally bad actions. Fear and disgust, however, did not lead to a reduction in the extremity of moral judgements for morally bad actions. This is further evidence for the particular influence of anger on control and moral judgement. These data appear in Appendix N.

General Discussion

In this chapter I have presented evidence inconsistent with the asymmetry hypothesis. Study 6 and 7 found that when the same uncontrollable impulse compels action, the extremity of moral judgement decreases for both morally good and bad actions. Study 8 found that uncontrollable anger reduces judgements of moral goodness, suggesting that it is not moral valence that is responsible for the asymmetry effect, but that participants view uncontrollable anger, but not uncontrollable sympathy, as reducing an agent’s control over their actions.

Pizarro et al. (2003) proposed that the asymmetry effect can be explained by the meta-desire hypothesis. According to the authors, second-order desires are an individual’s acceptance or rejection of a desire or impulse. For example, for a drug addict “there is a second-order desire (not to be an addict) and a first-order desire (a compulsion to continue to take drugs)” (Pizarro et al., p. 267). The meta-desire hypothesis states that the asymmetry effect is due to assumptions observers make about these second-order desires (or meta-desires). If the results from Studies 6-8 can be accounted for by the meta-desire hypothesis, then the case against the asymmetry hypothesis is less compelling, as an argument can be made that although Studies 6-8 were inconsistent with the findings from Pizarro et al. (2003), the results of Studies 6-8 studies can be explained away as special cases by the meta-desire hypothesis.

To assess blame for an uncontrollable action the authors argue it is reasonable to consider whether an individual endorses the actions they have been compelled to commit, and that the naïve observer has expectations about an agent’s second-order
desires. For an individual impulsively committing a morally bad action an observer has
the expectation that the individual has a second-order desire not to commit the morally
bad action. Because the second-order desire is inconsistent with the first-order impulse
the amount of blame the individual receives is reduced, because an observer infers that
the individual really did not want to commit the morally bad action, and only did so
because of an impulse. On the other hand, when an individual impulsively commits a
morally good action, observers have the expectation that the individual also has a
second-order desire to perform the morally good action. And so the first-order impulse
to perform the morally good action, and the inferred second-order desire to also perform
the morally good action are the same, and so discounting for a loss of control is not
needed.

However, the meta-desire account of the asymmetry effect is not without
problems (for a detailed analysis of the meta-desire hypothesis and implications for
Studies 6-8 see Appendix O). From Pizarro et al.’s description it is not clear what a
meta-desire is, and this lack of specificity makes interpreting the findings of Studies 6-8
in terms of the meta-desire hypothesis difficult. For example, on the same page (p. 267)
the authors describe a meta-desire at one point as “the true desire of the agent” and at
another point as “an individual’s higher-order acceptance or rejection of a desire or an
impulse”. These ideas are similar but lead to conflicting predictions about what the
agent’s true desire is, or whether an agent’s desires are consistent with the impulses they
have in Studies 6-8.

Furthermore, the empirical evidence for the meta-desire hypothesis is
questionable. In Study 2 of Pizarro et al. (2003) meta-desire information was
experimentally manipulated and it was concluded that the evidence was consistent with
the meta-desire account. However, a closer examination shows that the data from this
study is in fact inconsistent with a meta-desire account of the asymmetry effect
(Appendix O). A second study assessed the meta-desire hypothesis, and as well as
assessing participants praise and blame judgement for good and bad actions, Pizarro et
al. assessed participant’s meta-desire assumptions. It was found that meta-desires
mediated the asymmetry effect. However an alternative interpretation that fits equally
with the data, is that rather than inferred meta-desires influencing whether control
information is used to discount praise or blame, it is in fact assessments of control that
influence participants’ inferences of meta-desires. That is, when participants inferred that an agent acted deliberately they were more likely to say that the agent really wanted to perform the action, than if the agent’s actions were uncontrolled. It just happens that inferred meta-desires mediate the asymmetry effect because participants treat uncontrollable anger as a loss of control, while they do not treat uncontrollable sympathy as a reduction in control.

A related asymmetry
This chapter has focussed on the asymmetry hypothesis; however, two other studies have examined a similar asymmetry for praise and blame judgements of intentional actions. Perceived control is related to judgements of intentionality, as the loss of enough control is likely to result in an action being judged unintentional.

There is evidence that intentionality intensifies blame (Lagnado & Channon, 2008). Malle and Bennett (2002) argue that intentionality increases blame to a greater degree than intentionality increases praise. The authors conducted a study that manipulated whether the outcome of an action was morally good or bad, and whether the action was performed intentionally or unintentionally, and found that intentional actions compared with unintentional actions intensified blame to a greater degree than praise.

The evidence, however, is not compelling, as the study was conducted using a single vignette that described an odd situation. In the vignette an agent learns that a work colleague either loves or hates receiving phone calls at home, and the agent decides to call them at home at 6am the next day. The next day the agent either intentionally calls their work colleague or tries to call a different number but accidentally calls the work colleague. Quite apart from their moral culpability, it is hard to know what to make of a person who intended, but did not mean to call someone (and then called them anyway).

To examine whether the results of Malle and Bennett (2002) were due to the stimuli used, Ohtsubo (2007) ran a conceptual replication using a greater number of stimuli, and also found that blame was intensified by intentionality to a greater degree than was praise. However this study also features ambiguous vignettes. For example, “Your coworker, person U, opened the window of an air-conditioned room”.

95
Participants are then told that the action resulted in either a positive outcome (“You had been freezing in the low temperature, and you soon felt more comfortable”), or a negative outcome (“The temperature of the room rose and you got sick”). Finally, participants are told that the behaviour was performed intentionally or the behaviour was not performed intentionally. However, it is not clear whether this means the initial action the agent performs is intentional (opening the window) or whether opening the window to achieve a positive or negative outcome is intentional (opening the window to make the agent sick). And so the evidence for a related asymmetry effect of intentionality is not compelling.

The claims made by the studies considered in this chapter (Malle & Bennett, 2002; Ohtsubo, 2007; Pizarro et al., 2003) are about judgements of praise and blame. At least in the case of Pizarro et al. (2003), it is implied that these claims generalize to judgements of moral goodness, and in this chapter I treated judgements of praise and goodness as synonymous. However there are reasons to suspect that these potential asymmetries could influence judgements of praise and blame differently from judgements of moral good and bad. Considering evidence from negative actions, Cushman (2008) found that blame was more dependent on outcomes, and judgements of wrongness were more dependent on mental states of the person performing the action. So for both the control and intentionality asymmetry effects, one might predict the kind of judgement being made (praise/blame or good/bad) could be more or less sensitive to an asymmetry. For example, an asymmetry in the way control information resulting from an internal impulse is used between judgements of good or bad may be more likely to manifest, than between praise and blame, because judgements of good and bad are more dependent on the influence of mental states.

**Conclusion**

I have presented evidence for an alternative interpretation of the asymmetry effect. Initially I hypothesized that sympathy as a reason for performing an action could increase how morally good the action was considered to be, and that this would mask the reduction in praise for loss of control. However, a second hypothesis that anger and sympathy differ in how they are perceived to reduce control was supported by empirical evidence from Studies 6-8. The evidence suggests that the asymmetry effect disappears
when actions are performed with similar amounts of control for morally good and bad actions. I also argued that the meta-desire hypothesis was not able to account for the asymmetry effect. These results indicate that personal control is an important part of making goodness judgements of actions.
Chapter 6

General discussion

Research on moral judgement has focused almost exclusively on judgements of transgressions and dilemmas, addressing morally good actions, if at all, only as an afterthought. There is compelling evidence, however, that judgements of transgressions and goodness are made in different ways (Janoff-Bulman, Sheikh and Hepp, 2009; Takahashi et al., 2008; Wiltermuth, Monin and Chow, 2010), which suggests that claims about moral judgement based solely on studies of transgressions are problematic. This thesis addressed this notable gap in the literature by examining the concept of a morally good action, and investigating how judgements of moral goodness are made.

I began by considering what constitutes a morally good action. Study 1 examined the folk concept of moral goodness, broadly construed. Generally, participants identified that good actions are characterised by particular outcomes and motives. Although outcomes of actions have been the primary focus of the broader moral judgement literature, participants in fact mentioned the importance of motives more often than any other factor. Previous research has also shown that motives influence goodness judgements (Baird & Astington, 2004; Janoff-Bulman, Sheikh, & Hepp, 2009), but there has been almost no discussion, and even less data, on what specific motives are the “right” kinds of motives for performing morally good actions.

In Study 1, participants identified that being motivated by the interests of others, acting on principle, feeling sympathy, and not being motivated by self interest, are central to what it means for an action to be morally good. Since people often lack insight into what factors influence their judgements it was necessary to experimentally verify that the motives identified by participants, do in fact influence goodness judgements of actions. Studies 2-5 found that these motives do, empirically, influence goodness judgements, and together with Study 1, provide evidence that a person’s motive or reason for performing a good action influences perceptions of the goodness of their actions, and that certain motives and reasons are central to what it means for an action to be morally good.

Another important finding from Studies 2-5 was that the influence of motive on
goodness judgements of actions was partially mediated by perceptions of the actor’s character. That is, a positive action performed out of sympathy (for example) was judged morally good in part because the person performing it was judged to be morally good. Supporting this interpretation, previous research has found that motive is an indicator of character (Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002), and that perceptions of a person’s character influence moral evaluations of their actions (Alicke, 1992, 2000; Nadler & McDonnell, 2011).

According to another more speculative account that is challenging to the current conception of moral judgements of actions, it may be that when people evaluate actions they are essentially evaluating the person who performed the action (Pizarro & Tannenbaum, 2012). According to this hypothesis a good action is the sort of action that a good person performs, and the response to the question “how morally good was it that a person did X”, is essentially an evaluation of whether the action is indicative of the actions of a good person. It may be, that rather than character judgements indirectly mediating the relationship between motive and goodness judgements, that motives reveal the good or bad character of a person, which is then reflected by evaluations of that person’s actions. A similar distinction between action and person as targets of moral evaluation can be found in the distinction between consequentialist and deontological ethics, which evaluate actions, and virtue ethics where the target of moral evaluation is a person. However, when interpreting the mediating role of character, as with all mediation analyses, there remains the question of causality: the fact that character mediates judgments is necessary but not sufficient for character to have caused those judgments. What is needed, and a good subject for subsequent research, is a study in which motivation and character are manipulated independently.

The insight that motive and reasons influence goodness judgements of actions was then used to interpret the asymmetry effect – the fact that judgements of wrongdoing are discounted when an actor lacks control, but goodness judgements of positive actions are not (Pizarro, Uhlman, & Salovey, 2003). The asymmetry effect is one of the most important findings in the literature on goodness judgements of actions, as it calls into question the importance of an agent’s control over their actions. It was hypothesized that praise was not discounted in the uncontrollable sympathy condition,

---

31 Pizarro & Tannenbaum (2012) argue specifically for blame judgements.
because the presence of sympathy increases the moral goodness of the action (as found in Studies 3-5), masking the fact that the uncontrollable nature of the action would otherwise have reduced praise. However, in the process of designing a study to examine this hypothesis, a simpler explanation for the asymmetry effect became apparent: uncontrollable anger (as a motivation for negative actions) implies a greater loss of control than uncontrollable sympathy (as a motivation for positive actions). This alternative account was examined experimentally in Studies 6-8, and it was determined that the asymmetry effect and its proposed explanation, the meta-desire hypothesis, do not appear to hold up when control is equated between good and bad actions.

As noted, unlike the blame literature (Alicke, 2000; Guglielmo, Monroe, & Malle, 2009; Shaver, 1985) there are currently no empirically derived models describing how factors relevant to goodness judgements are used. Candidate factors for goodness judgements informed by blame models include: a positive outcome, intention, motivation, a causal relationship between the agent and the outcome, and sufficient control by the agent over their thoughts and actions. Blame models suggest that at least some of these factors may be necessary for blame, and that there will be strong interactions between the influence of factors (Guglielmo, Monroe, & Malle, 2009; Shaver, 1985). For example, the influence of intentionality on blame should be strongly influenced by whether a negative outcome has occurred (Guglielmo, Monroe, & Malle, 2009). However, there is also evidence that some factors are independent. Cushman (2008) examined blame and wrongness judgements and found the presence of a negative outcome did not strongly influence the effect of intention on wrongness judgements, and that the influence of intention and outcome information was additive. If these findings generalize to praise and goodness judgements, then we may find, as was hinted at in Study 2, that motives can influence goodness judgements independently of a particular good outcome.

It should be noted that, although goodness and praise were not empirically distinguished, some research suggests that they should be. Cushman (2008) also found that blame judgements were more sensitive to information about outcomes, while wrongness judgements were more sensitive to the intent of the person performing the action. This suggests that there could be important differences between praise and goodness judgements in the use of particular kinds of information, and that an attempt
should be made to clarify differences in sensitivity of praise and goodness judgements to different kinds of information. Of particular relevance to this thesis, it may be that goodness judgements, more so than praise, are sensitive to information about motives.

Another factor that may have important consequences for goodness judgements is how an agent feels when deciding to perform, or while performing, a morally good action. I argued that sympathy is the right kind of reason for performing a good action, and when present it increases how morally good an action is evaluated to be. However, it may be that instead of (or in addition to) being the right kind of reason, sympathy also indicates how an agent feels when they perform a good action, and that feeling the right or wrong way when performing a good action will influence goodness judgements of that action. For example, if someone donates money to disaster relief because they have a principled belief that it is important to give to those in need, but they are not feeling particularly compassionate towards those who are suffering, this may influence how morally good donating money will be judged. Analogously, in the current studies, it is not clear whether the goodness of an action (and of the agent) was reduced because the agent denied acting upon a particular good reason, or because the agent denied feeling the right way, which may be an important aspect of performing a good action. Future research could clarify the relationship between how a person feels when performing a good action, and the observer’s goodness judgements of the agent’s actions.

In this thesis a number of motives associated with good actions were investigated, however one motive that appears in the literature (Janoff-Bulman, Sheikh, & Hepp, 2009), but was not identified in Study 1 as a motive, is duty. Duty as discussed by Janoff-Bulman et al. (2009) is characterised as performing a good action because a person thinks that they should perform the action. Duty has also been described as indicating that a person ought to perform, or that a person has an obligation to perform an action (Brandt, 1964). Duty plays an important part in philosophical discussions of goodness, and in fact for Kant, an action must be motivated by duty to have moral worth (Kant, 1785/2002). Duty differs from the other motives considered here – other interest, principle and sympathy – in terms of the implications of obligation, and it is interesting to note that, in Study 1 some participants explicitly stated that that a good action is one that a person is not obligated to perform. These responses were coded into the category no obligation, but perhaps these responses could be thought of as indicating that
morally good actions are not motivated by an attempt to fulfil one’s duty. This tension between the importance of duty in the philosophical and psychological literatures, and the reports of participants in Study 1 that good actions are not obligatory, could be examined in further research into the role of duty as a motive for good actions, and its relation to the other motives identified.

Janoff-Bulman et al. (2009) also make a potentially important distinction between two kinds of morally good actions: behaving in a positive way (e.g., volunteering to clean garbage off a beach), and resisting the temptation to behave in a negative way (e.g., avoiding gambling when a person has a financial responsibility to their family). Janoff-Bulman et al. found that the influence of duty on goodness judgements differed depending on which type of good action an agent performed. In Study 2 both kinds of good were used to examine the influence of motives, however no difference in goodness judgements for actions motivated by self interest or principle were observed.

Although an interpretation based on reasons was not needed to explain the asymmetry effect found by Pizzaro et al. (2003), the influence of motives and reasons can be used to explain other findings in the literature. For example, Woolfolk, Doris, and Darly, (2006) found that when a person was compelled to transgress (e.g., ordered to kill another person at gunpoint, or was administered a “compliance drug” and forced to kill another person), the more the agent “identified” with the action, the more morally responsible participants held them for killing the person. “Identifying” with an action, according to the authors, means that a person embraces or wholeheartedly performs the action. In Woolfolk et al., an agent either identified with killing another person (e.g., had decided earlier that they were going to kill the person), or they did not identify (e.g., did not have a motive to kill them). The authors found that agents who identified with killing the person were judged to be more morally responsible for their actions. An alternative explanation to the “identification” account, however, is that in the identification condition the agent has a motive to kill the person, and in the low identification case they do not have a motive. It may simply be that the presence of a motive to kill another person increases judgements of moral responsibility because a desire to kill or harm another person is closely associated with what it means for an action to be wrong, and so the agent is deemed more morally responsible for the
person’s death.

Another way that motives may influence moral judgements is subtler. Motives are left unstated in most studies and participants must infer them from the agent’s actions or other information. In many studies, the participant is asked to imagine themselves in a dilemma, and then to judge what the right course of action is (e.g., Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Schnall, Haidt, Clore, & Jordan, 2008). In other studies, however, the participant is told that another person has a moral dilemma, or has performed a potential transgression, and that the participant should evaluate the other person’s actions (e.g., Cushman, Young, & Hauser, 2006; Haidt, Koller, & Dias, 1993). It is likely that the motive attributed to the agent in a particular scenario, will be influenced by whether the agent is described as being the participant or another person. Evidence suggests that people have a tendency to portray themselves positively (Sedikides & Strube, 1997), and to explain their own actions using reason explanations (e.g., other interest and principle), while more frequently explaining the behaviours of others using causal history explanations (e.g., sympathy; Malle, Knobe, & Nelson, 2007). The tendency to attribute different motives to self than others may result in the same situation being evaluated differently depending on whether the agent in the situation is the participant or another person.\footnote{Furthermore Alicke (2000) argues that people have a blame validation process, whereby observers are more likely to favour blame versus no blame explanations of events. Consistent with the blame validation approach people may seek to infer motives that are consistent with a blameworthy action. However given what we know about the self serving bias (Campbell & Sedikides, 1999) it seems unlikely that people will search for motives that are consistent with themselves being blameworthy.}

An important methodological limitation of the studies presented in this thesis is the way in which aspects of the concept of a good action, and particularly motives and reasons, were identified and distinguished. Aspects of moral goodness and motives were initially identified from participants’ responses in Study 1. This method appears to be sufficient for identifying broad categories, but finer-grained distinctions between motives was difficult and was mostly speculative. This became more apparent in the difficulties distinguishing self interest and neutral motives, and with overlap between motives of other interest and principle. As discussed, clearly articulating the breadth of ordinary language concepts is a difficult thing to do. This research presented in this thesis is an initial approach to moral goodness and motives, but a more detailed and
sophisticated approach to identifying and distinguishing motives and reasons is an important next step in understanding the motives associated with the good.

A further limitation is the homogeneous demographics of the sample. Participants were all university students of a similar age and background attending a New Zealand university. However, as noted in the Introduction, although it is important to generalize these findings to other groups, there is already evidence for broadly similar moral intuitions across gender, education, politics, and religion (Banerjee, Guebner & Hauser, 2010). Furthermore, understanding the goodness judgements of the current and similar samples is important as they are typical of most previous research in the area, and a primary goal of this thesis is to inform that literature. Nevertheless, replication within and across diverse samples will ultimately be necessary for any complete model of moral judgment.

Another limitation relevant to issues of generalizability is that most of the current studies used vignettes in their methodology. Although this methodology is standard practice in the literature on moral judgment, using only vignettes, especially a relatively small number, increases the chances that the findings are methodology-specific. To increase confidence in the results further research should use a greater number and variety of vignettes as well as using other non-vignette methodologies, for example, the influence of inferred and reported motives on goodness judgements could be examined by having participants observe real people performing good actions in the lab.

In summary, the focus of research on moral judgement has been on moral transgressions and how people evaluate the moral goodness of actions is not well understood. This thesis took an initial exploratory look at the concept of a morally good action, and through a series of studies, has advanced understanding of the relation between motive, character, control and goodness judgements of actions.
References


Broeders, R., van den Bos, K., Müller, P. A., & Ham, J. (2011). Should I save or should I not kill? How people solve moral dilemmas depends on which rule is most accessible. *Journal of Experimental Social Psychology, 47*(5), 923-934.


Gazzaniga, M. S. (2000). Cerebral specialization and interhemispheric communication Does the corpus callosum enable the human condition?. *Brain, 123*(7), 1293-1326.


Appendix A

Study 1 open response

#_____________

In this next task we are interested in your thoughts on morality.

When you say that something is *morally good*, what do you mean? Please explain.

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________
Appendix B

Study 1 instruction for coders

When you say that something is *morally good*, what do you mean?

The participant’s response is to be categorized into the following categories. Responses may fall into all, some or none of the categories.

*Note: the category names are an indication of the category. They aren’t perfect descriptions and the category may be wider or narrower than the category name suggests.

**Action – character**

Think about whether the respondent is talking about an action that a person is performing or referring to an evaluation of a person. For example an action would be “when someone does something morally good they think about the consequences of their actions” a reference to character would be more like “a morally good person thinks about the consequences of their actions”

**Action**
Something someone does.

“something that you would do”
“behaving in a way”
“an act”

Participants may not explicitly talk about performing an action but if their writing indicates that they are talking about ways in which something someone does can be morally good then this is classified as Action.

**Character**

Classify as action if the participant writes about ways in which a person can be morally good or characteristics, traits or qualities a morally good person has.

“cant have a morally good axe murderer who felt sorry, not can you have a morally good benefactor who has hidden agendas”
“something that portrays a good character “ – talking about character reflected by an action
“a moral person is someone who”
“someone who”
“a morally good person is”
“being morally good”

Not to be confused with if someone has individual morals or beliefs – this would be
categorised as principle

**Synonyms**

The category synonym is used when a participant describes what a morally good action is using a synonym for morally good. These statements are often of the form: Something that is morally good is …

good/ethical/just/right thing to do
“It’s the right thing to do”

**Principle**

Principles are morals, rules, codes, values etc that people can follow or adhere to.
“a persons morals”
“have a set of good standards”
“based on good values” #someone who has good values would be character
“depending on moral codes”
“doesn’t contradict my principles”
“consistent with your morals”
“it[the action] fits within our morals”
“good standards”
“guidelines”
“upholding justice”

**Other interest**

Participant mentions that something is morally good when a person has:
- Good intentions
- Considers others needs/people /care for others good intentions
- Consider how something effects/benefits others
- Motivated to help
- Other directed thoughts
- Holding others in good regard – respect and their consequences

“intentions are good”
“respect for others”
“done purely to help others”

*Note: this category is to be used to indicate that the participant has mentioned a way in which a person thinks or is motivated to act towards another person or thing. If the participant is just talking about the consequences of an action e.g. “people benefit” then that would be categorized as good for others. Good for others is a category for outcomes

**Sympathy**

---

33 The original category name when coders were coding the data was
34 The original category name when coders were coding the data was “Empathy”
Participant mentions empathy, compassion or a way someone should feel about another person or thing
“have empathy and compassion”

**Benefit others**

Use this category if participant mentions some way in which the action benefits others. Benefit others / outcome is in the interest of others / treat others well / help others

Is good to others?

“must benefit a person or society positively”
“best outcome for everyone”
“help others”
Gives an example of a behaviour that helps people
“cause fewer people hurt or anger”

**No motive to harm**

Use this category if a participant indicates that a person must not have the motivation to harm others

**No harm**

Like the category good of others this category is used to indicate outcomes of actions. In this case the outcome doesn’t impact negatively on others. Participants often indicate that the action doesn’t harm others at all, a few may say that there is more good then harm or harm is lessened.

“something that doesn’t negatively affect others”
“... without negatively impacting anyone or anything”
“doesn’t negatively affect other’s well-being”
“that it doesn’t do anyone any harm”
“doesn’t restrict the rights of others”

Give examples of harmful behaviour
Included here and in good of others, would be someone saying that it is something that does more good than bad
It’s interesting how frequently the idea of not harming people in any way appears

**No self interest**

Participant indicates that a person must have no motive to benefit themselves

“selfless”
“no ulterior motives”
“wants no gain”

---

35 The original category name when coders were coding the data was “Good for others”
36 The original category name when coders were coding the data was, No motive to benefit self
“not just your own motives”
“Give thought to not benefiting self”
“There is no motive to benefit yourself”

**No benefit to self**

A person does not benefit from the action.

“cares less about self than others”
“action doesn’t benefit yourself”
“not only in your own personal best interest”
“an act that may help others even at the disadvantage of the person performing it”

**Justification**

These are responses that indicate how morality is justified. If the participant mentions rules or principles this will also fall under the principle category.

**Society**

“something which is considered a good action within your society”
“fills the societal beliefs of what is right”
“sometimes behaving in a way which is socially normal”
“It’s morals that are taught and a life lesson”
“something accepted by the large population”

**Self**

“your own sense of what is right”
“… but I think your own judgement or sense of what is right”
“person belief of what is good or bad”

**Religion**

“I think without god there is no justification”

**Subjective**

People may mention that morals are either subjective or objective.

If a participant explicitly indicates that morality is subjective.

**Objective**

If participant explicitly indicates that morality is objective.

Universal rules

---

37 Original category name when coders were coding was Don’t benefit self
Feel good don’t regret

Participant mentions that something is morally good, or is likely to be morally good if it makes you feel a certain way

“makes one feel good”
“something that you would do without any regrets and be totally happy with after”
“something that you won’t feel poorly for doing”

Feeling

A sense or feeling indicates what is right and wrong

“using your own inner ‘sense’ of what is the correct way to behave.
“an innate sense”
“gut feeling”
“sense”

Strong belief

Moral beliefs are strong

“When saying morally good or bad you have strong feelings”
“something you strongly believe in”

No obligation

Doing something that is morally good does not mean you are just fulfilling an obligation

“not something you are bound to do by the law”

Training examples

Something that is a positive act that causes no harm to anybody else. A “morally good” act is usually selfless and has positive benefits.

If a person acts towards a principle they have by society then that’s moral. They think about others and the thing they do is good for others.

A good person thinks about other people. You know something is morally good because you feel good about it. A good action wouldn’t harm others.
Someone who has can empathise with others. I think that you get your morals from god, and your family. You do something that you don’t regret later.

Not gaining yourself and hurting others.
If someone has good morals they take into account others. And they don’t just do it because they are expected to do.
Appendix C

Study 2 vignettes and questions

Instructions:
In this task we are interested in how people evaluate others' actions. To do this we will present you with several situations and ask you what you think about the actions of the people in each situation. We are interested in how morally good you think each action is. Each action described may be good, bad or neither.

Each question will be accompanied by a scale. To evaluate an action please press the number on the keyboard that corresponds to the scale.

Example Question:
Rebecca didn't have enough money for lunch so Jon gave her $10 so she could buy lunch.

1(Extremely morally bad) – 5(Morally neutral) – 9(Extremely morally good)

How morally good or bad was it that Jon gave Rebecca $10?

Please read each question carefully.

When you are ready to begin press the spacebar.

The following questions were asked after participants read each vignette.

• How morally good or bad was it that X? 1(Extremely morally bad, 5(morally neutral), 9(Extremely morally good)

• Do you think that Y is mainly a good or bad person? 1(Mainly a bad person, 9(Mainly a good person)

• Do you think that Y has good moral standards? 1(Not at all), 9(Completely)

Vignette 1

Neutral: Georgia walked to work today because it looked like a nice day and she didn’t want to drive her car.

Other interest: Georgia walked to work today because she wanted to help the environment by not driving her car.

Principle: Georgia walked to work today because she has a moral belief that protecting the environment is important so she didn’t drive her car to work.
Self interest: Georgia walked to work today instead of driving her car because she wanted to lose weight.

Self/other interest: Georgia walked to work today instead of driving her car because she wanted to help the environment and she also wanted to lose weight.

Questions:
How morally good or bad was it that Georgia walked to work?
Do you think that Georgia is mainly a good or bad person?
Do you think that Georgia has good moral standards?

Vignette 2

Neutral: Daniel decided to work late because he is trying to impress his work colleagues.

Other interest: Daniel decided to work late because he wanted to help his work colleagues by getting his work done.

Principle: Daniel decided to work late because he has a moral belief that working hard and getting your work done is important.

Self interest: Daniel decided to work late because he thought it would improve his chances of getting a promotion.

Self/Other interest: Daniel decided to work late because he wanted to help his work colleagues by getting his work done and he also thought it would improve his chances of getting a promotion.

Questions:
How morally good or bad was it that Daniel decided to work late?
Do you think that Daniel is mainly a good or bad person?
Do you think that Daniel has good moral standards?

Vignette 3

Neutral: Emma is sick and decides not to go to work today because she isn’t feeling very well.

Other interest: Emma is sick and decides not to go to work today because she doesn’t want to make her work colleagues sick.

Principle: Emma is sick and decides not to go to work today because she has a moral belief that not harming other people is important and she doesn’t want to make her work colleagues sick.

38 This falls under the description of self interest
**Self interest:** Emma is sick and decides not to go to work today because she wants sympathy from her work colleagues when she goes back to work.

**Self/Other interest:** Emma is sick and decides not to go to work today because she doesn’t want to make her work colleagues sick and she also wants sympathy from her work colleagues when she goes back to work.

Questions:
How morally good or bad was it that Emma decided not to go to work?
Do you think that Emma is mainly a good or bad person?
Do you think that Emma has good moral standards?

Vignette 4

**Neutral:** Jacob’s wife is out shopping and he is at home by himself so he decides to wash the dishes because he is bored and thinks that he may as well get them done.

**Other interest:** Jacob’s wife is out shopping and he is at home by himself, so he decides to wash the dishes because he wants to make the house nice for his wife when she gets home.

**Principle:** Jacob’s wife is out shopping and he is at home by himself so he decides to wash the dishes because he has a moral belief that helping others is important so he washes the dishes to make the house nice for his wife when she gets home.

**Self interest:** Jacob’s wife is out shopping and he is at home by himself so he decides to wash the dishes because he wants his wife to be grateful.

**Self/Other interest:** Jacob’s wife is out shopping and he is at home by himself so he decides to wash the dishes because he wants to make the house nice for his wife when she gets home and he also wants his wife to be grateful.

Questions:
How morally good or bad was it that Jacob washed the dishes?
Do you think that Jacob is mainly a good or bad person?
Do you think that Jacob has good moral standards?

Vignette 5

**Neutral:** Olivia tells her family that she has decided to eat more healthily because she wants to look good.

**Other interest:** Olivia tells her family that she has decided to eat more healthily because she wants to be healthier so she can do a better job looking after her family.

**Principle:** Olivia tells her family that she has decided to eat more healthily because she has a moral belief that you have a responsibility to look after your family and if she is healthier she can do a better job looking after them.
Self interest: Olivia tells her family that she has decided to eat more healthily because she wants her friends to think she looks really good.

Self/Other interest: Olivia tells her family that she has decided to eat more healthily because she wants to be healthier so she can do a better job looking after her family and she also wants her friends to think she looks really good.

Questions:
How morally good or bad was it that Olivia decided to eat more healthily?
Do you think that Olivia is mainly a good or bad person?
Do you think that Olivia has good moral standards?

Vignette 6

Neutral: It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he is tired.

Other interest: It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he doesn’t want to annoy his neighbours by playing music late at night.

Principle: It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he has a moral belief that you shouldn’t behave selfishly so he turns off the music and goes to be bed so he doesn’t annoy her neighbours by playing music late at night.

Self interest: It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he doesn’t want to annoy the girl next door who he is trying to impress.

Self/Other interest: It is very late when James gets home from work in the evening. He sits in the lounge for a while listening to music and thinking about his day. James decides to turn off the music and go to bed because he doesn’t want to annoy his neighbours by playing music late at night and he also doesn’t want to annoy the girl next door who he is trying to impress.

Questions:
How morally good or bad was it that James turned off the music and went to bed?
Do you think that James is mainly a good or bad person?
Do you think that James has good moral standards?

Vignette 7
Neutral: Hannah sees a friend working on an assignment. Hannah talks to her friend for a while and then leaves to eat some food because she was starting to feel hungry.

Other interest: Hannah sees a friend working on an assignment. Hannah talks to her friend for a while and then leaves because she doesn’t want to distract her friend from working.

Principle: Hannah sees a friend working on an assignment. Hannah talks to her friend for a while and then leaves because she has a moral belief that helping people is important and she wants to help her friend finish her assignment by not distracting her from her work.

Self interest: Hannah sees a friend working on an assignment. Hannah talks to her friend for a while and then leaves because she wants to work on her own assignment.

Self/Other interest: Hannah sees a friend working on an assignment. Hannah talks to her friend for a while and then leaves because she doesn’t want to annoy her friend by distracting her and she also wants to work on her own assignment.

Questions:
How morally good or bad was it that Hannah left her friend?
Do you think that Hannah is mainly a good or bad person?
Do you think that Hannah has good moral standards?

Vignette 8

Neutral: Thomas is being considered for a promotion at work and so is his friend. Thomas decides to withdraw from consideration for the promotion because he doesn’t want the extra work.

Other interest: Thomas is being considered for a promotion at work and so is his friend. Thomas decides to withdraw from consideration for the promotion because he thinks that his friend deserves the promotion more than him and that his friend should get the job.

Principle: Thomas is being considered for a promotion at work and so is his friend. Thomas decides to withdraw from consideration for the promotion because he has a moral belief that you shouldn’t be selfish and he thinks that his friend deserves to get the job more than him.

Self interest: Thomas is being considered for a promotion at work and so is his friend. Thomas decides to withdraw from consideration for the promotion because he doesn’t think the job pays enough and thinks that he will make more money working somewhere else.

Self/Other interest: Thomas is being considered for a promotion at work and so is his friend. Thomas decides to withdraw from consideration for the promotion because he thinks that his friend deserves the promotion more than him and that his friend should
get the job he also doesn’t think the job pays enough and thinks that he will make more money working somewhere else.

Questions:
How morally good or bad was it that Thomas withdrew from consideration?
Do you think that Thomas is mainly a good or bad person?
Do you think that Thomas has good moral standards?
Appendix D

Study 3 vignettes and questions

Initial instructions:

In this task we are interested in how people evaluate others' actions. To do this we will present you with several situations and ask you what you think about the actions of the people in each situation.

Each question will be accompanied by a scale. To evaluate an action please press the number on the keyboard that corresponds to the scale.

Example Question:
Matthias gave Jon $10 for lunch because Jon didn't have enough money to pay for lunch.

How much praiseworthy was it that Matthias gave Jon $10?
1(Not at all praiseworthy) – 7(Extremely praiseworthy)

For each situation you will be asked several questions and some of them are quite similar. Please read each question carefully.

When you are ready to begin press the spacebar.

Second sympathy condition instructions:

Next you will be presented with the same situations again. This time however some of the reasons for why the people behaved in the way they did have CHANGED.

PLEASE READ EACH QUESTION CAREFULLY BEFORE YOU ANSWER. Your response may be the same or different from the first time you read about the situations.

Participants were asked the following questions after each vignette.

**How praiseworthy was it that x did y?**
1(Not at all praiseworthy) – 7(Extremely praiseworthy).

**How morally good was it that x did y?**
1(Not at all morally good) – 7(Extremely morally good).

**How positively should x be evaluated for y?**
1(Not at all positively) – 7(Extrememly positively).

Volunteered vignettes
Vignette 1

*Sympathy:* Jackie donated $200 to help the victims of a flood. When asked why she donated the money she replied, “I felt enormous sympathy for the victims and it was the right thing to do”.

*No sympathy:* Jackie donated $200 to help the victims of a flood. When asked why she donated the money she replied, “It was the right thing to do”.

Questions:
How praiseworthy was it that Jackie donated $200?
How morally good was it that Jackie donated $200?
How positively should Jackie be evaluated for donating $200?

Vignette 2

*Sympathy:* Nicholas works as a builder. He heard about a family who lost their house in a fire. Nicholas volunteered to help the family rebuild their house without charging them any money. When asked why he volunteered to help he replied, “I felt enormous sympathy for the family and it was the right thing to do”.

*No sympathy:* Nicholas works as a builder. He heard about a family who lost their house in a fire. Nicholas volunteered to help the family rebuild their house without charging them any money. When asked why he volunteered to help he replied, “It was the right thing to do”.

Questions:
How praiseworthy was it that Nicholas volunteered to help the family rebuild their home?
How morally good was it that Nicholas volunteered to help the family rebuild their home?
How positively should Nicholas be evaluated for helping the family rebuild their home?

Vignette 3

*Sympathy:* While waiting in line at a checkout to pay for her groceries Sally notices that the lady in front of her doesn’t have enough money to pay for her groceries. Sally gives the lady $50 to pay for her groceries. When asked why gave the lady money for her groceries she replied, “I felt enormous sympathy for her and it was the right thing to do”.

*No sympathy:* While waiting in line at a checkout to pay for her groceries Sally notices that the lady in front of her doesn’t have enough money to pay for her groceries. Sally gives the lady $50 to pay for her groceries. When asked why gave the lady money for her groceries she replied, “It was the right thing to do”.

129
Questions:
How praiseworthy was it that Sally gave the lady $50 for her groceries?
How morally good was it that Sally gave the lady $50 for her groceries?
How positively should Sally be evaluated for giving the lady $50 for her groceries?

Questioned vignettes

Initial instructions:

In the next section we will be asking you again to evaluate how morally good some actions are.

As before each question will be accompanied by a scale. To evaluate an action please press the number on the keyboard that corresponds to the scale.

For each situation you will be asked several questions\(^{39}\) and some of them are quite similar. Please read each question carefully.

Second sympathy condition instructions:

Next you will be presented with the same situations again. This time however some of the reasons for why the people behaved in the way they did have CHANGED.

PLEASE READ EACH QUESTION CAREFULLY BEFORE YOU ANSWER. Your response may be the same or different from the first time you read about the situations.

Vignette 1

*Sympathy*: Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “I feel enormous sympathy for him and I think it’s the right thing to do”.

*No sympathy*: Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “Not really, I just think it’s the right thing to do”.

Questions:
How praiseworthy is it that Matthew visits his friend in hospital each day?
How morally good is it that Matthew visits his friend in hospital each day?

\(^{39}\) This should have read situations.
How positively should Matthew be evaluated for visiting his friend in hospital each day?

Vignette 2

*Sympathy:* Sharon read that the government has recently cut funding for home care for the elderly. She decides to volunteer to work 4 hours a week helping elderly people with jobs around the house that they have difficulty doing themselves. When asked whether she volunteered because she felt sympathy for the elderly people she replied, “I felt enormous sympathy for them and I thought it was the right thing to do”.

*No sympathy:* Sharon read that the government has recently cut funding for home care for the elderly. She decides to volunteer to work 4 hours a week helping elderly people with jobs around the house that they have difficulty doing themselves. When asked whether she volunteered because she felt sympathy for the elderly people she replied, “Not really, I just thought it was the right thing to do”.

Questions:
- How praiseworthy was it that Sharon volunteered to help the elderly people?
- How morally good was it that Sharon volunteered to help the elderly people?
- How positively should Sharon be evaluated for volunteering to help the elderly people?

Vignette 3

*Sympathy:* Jon just watched a news story about people getting sick because they don’t have access to clean drinking water. Jon decides to donate $100 to an organization that helps provide clean drinking water to those who don’t have access to it. When asked whether he donated the money because he felt sympathy for the people who didn’t have access to clean water he replied, “I felt enormous sympathy for them and I thought it was the right thing to do”.

*No sympathy:* Jon just watched a news story about people getting sick because they don’t have access to clean drinking water. Jon decides to donate $100 to an organization that helps provide clean drinking water to those who don’t have access to it. When asked whether he donated the money because he felt sympathy for the people who didn’t have access to clean water he replied, “Not really, I just thought it was the right thing to do”.

Questions:
- How praiseworthy was it that Jon donated $100 to help people without clean drinking water?
- How morally good as it that Jon donated $100 to help people without clean drinking water?
- How positively should Jon be evaluated for donating $100 to help people without clean drinking water?
Appendix E

Study 4a volunteered vignettes

Initial instructions

In this task we are interested in how people evaluate others' actions. To do this we will present you with several situations and ask you what you think about the actions of the people in each situation.

Each question will be accompanied by a scale. To evaluate an action please press the number on the response box that corresponds to the scale.

Example Question:
Matthias gave Jon $10 for lunch because Jon didn't have enough money to pay for lunch.

How much morally good was it that Matthias gave Jon $10?

1(Morally neutral) – 7(Extremely morally good)

Please read each question carefully. When you are ready to begin press the spacebar to begin.

Instructions 2

In a previous task you were presented with situations and asked to evaluate the actions of people in the situations. In this task you will be presented with the same situations you saw earlier. This time however some of the reasons for why the people behaved in they way they did have CHANGED.

PLEASE READ EACH QUESTION CAREFULLY BEFORE YOU ANSWER. Your response may be the same or different from the first time you read about the situations.

Please press any button on the response box to continue.

Question 1

Sympathy

Jackie donated $10 to help the victims of a flood. When asked why she donated the money she replied, “I felt a lot of sympathy for the victims and it was the right thing to do”.
How morally good was it that Jackie donated money?

1 (Morally neutral) – 7 (Extremely morally good)

No Sympathy

Jackie donated $10 to help the victims of a flood. When asked why she donated the money she replied, “It was the right thing to do”.

How morally good was it that Jackie donated money?

1 (Morally neutral) – 7 (Extremely morally good)

Question 2

Sympathy

Oliver works with Andrew. It’s lunch time and Andrew is very stressed. He has a lot of work to do and doesn’t have enough time to get lunch. Oliver knows that Andrew is stressed and working through lunch, so he buys him a sandwich so he can eat it at his desk. When asked why he bought Andrew a sandwich Oliver replied, “I felt a lot of sympathy for him and it was the right thing to do”.

How morally good was it that Oliver bought Andrew a sandwich?

1 (Morally neutral) – 7 (Extremely morally good)

No Sympathy

Oliver works with Andrew. It’s lunch time and Andrew is very stressed. He has a lot of work to do and doesn’t have enough time to get lunch. Oliver knows that Andrew is stressed and working through lunch, so he buys him a sandwich so he can eat it at his desk. When asked why he bought Andrew a sandwich Oliver replied, “It was the right thing to do”.

How morally good was it that Oliver bought Andrew a sandwich?

1 (Morally neutral) – 7 (Extremely morally good)

Sympathy

Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked why he visits his friend in the hospital Matthew replied, “I feel a lot of sympathy for him and it is the right thing to do”.

How morally good is it that Matthew visits his friend in hospital?
1 (Morally neutral) – 7 (Extremely morally good)

No Sympathy

Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked why he visits his friend in the hospital he replied, “It is the right thing to do”.

How morally good is it that Matthew visits his friend in hospital?

1 (Morally neutral) – 7 (Extremely morally good)
Appendix F
Study 4b questioned vignettes

Initial instructions

In the next section we will be asking you to evaluate how morally good some actions are.

As before each question will be accompanied by a scale. To evaluate an action please press the number on the keyboard that corresponds to the scale.

For each situation you will be asked several questions and some of them are quite similar. Please read each question carefully.

When you are ready to begin press the spacebar.

Instructions 2

Next you will be presented with the same situations again. This time however some of the reasons for why the people behaved in the way they did have CHANGED.

PLEASE READ EACH QUESTION CAREFULLY BEFORE YOU ANSWER. Your response may be the same or different from the first time you read about the situations.

Please press the spacebar to continue.

Vignette 1

Sympathy
Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “I feel enormous sympathy for him and I think it’s the right thing to do”.

No sympathy
Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Each day Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “Not really, I just think it’s the right thing to do”.

Questions:
How morally good is it that Matthew visits his friend in hospital each day?
1 (morally neutral) – 7 (extremely morally good)
Do you think that Matthew is mainly a good or bad person?
1 (mainly a bad person) – 7 (mainly a good person)

Do you think that Matthew has good moral standards?
1 (not at all positively) – 7 (extremely positively)

Vignette 2

Sympathy
Sharon read that the government has recently cut funding for home care for the elderly. She decides to volunteer to work 4 hours a week helping elderly people with jobs around the house that they have difficulty doing themselves. When asked whether she volunteered because she felt sympathy for the elderly people she replied, “I felt enormous sympathy for them and I thought it was the right thing to do”.

No sympathy
Sharon read that the government has recently cut funding for home care for the elderly. She decides to volunteer to work 4 hours a week helping elderly people with jobs around the house that they have difficulty doing themselves. When asked whether she volunteered because she felt sympathy for the elderly people she replied, “Not really, I just thought it was the right thing to do”.

Questions:
How praiseworthy was it that Sharon volunteered to help the elderly people?
1 (morally neutral) – 7 (extremely morally good)

How morally good was it that Sharon volunteered to help the elderly people?
1 (mainly a bad person) – 7 (mainly a good person)

How positively should Sharon be evaluated for volunteering to help the elderly people?
1 (not at all positively) – 7 (extremely positively)

Vignette 3

Sympathy
Jon just watched a news story about people getting sick because they don’t have access to clean drinking water. Jon decides to donate $100 to an organization that helps provide clean drinking water to those who don’t have access to it. When asked whether he donated the money because he felt sympathy for the people who didn’t have access to clean water he replied, “I felt enormous sympathy for them and I thought it was the right thing to do”.

No sympathy
Jon just watched a news story about people getting sick because they don’t have access to clean drinking water. Jon decides to donate $100 to an organization that helps provide clean drinking water to those who don’t have access to it. When asked whether
he donated the money because he felt sympathy for the people who didn’t have access to clean water he replied, “Not really, I just thought it was the right thing to do”.

Questions:
How praiseworthy was it that Jon donated $100 to help people without clean drinking water? 1 (morally neutral) – 7 (extremely morally good)

How morally good as it that Jon donated $100 to help people without clean drinking water? 1 (mainly a bad person) – 7 (mainly a good person)

How positively should Jon be evaluated for donating $100 to help people without clean drinking water? 1 (not at all positively) - 7 (extremely positively)
Appendix G

Study 5a volunteered vignettes

Initial instructions

In this task we are interested in how people evaluate others' actions. To do this we will present you with several situations and ask you what you think about the actions of the people in each situation.

Each question will be accompanied by a scale. To evaluate an action please press the number on the response box that corresponds to the scale.

Example Question:
Matthias gave Jon $10 for lunch because Jon didn't have enough money to pay for lunch.

How much morally good was it that Matthias gave Jon $10?
1(Morally neutral) – 7(Extremely morally good)

Please read each question carefully. When you are ready to begin press the spacebar to begin.

Instructions 2

In a previous task you were presented with situations and asked to evaluate the actions of people in the situations. In this task you will be presented with the same situations you saw earlier. This time however some of the reasons for why the people behaved in the way they did have CHANGED.

PLEASE READ EACH QUESTION CAREFULLY BEFORE YOU ANSWER. Your response may be the same or different from the first time you read about the situations.

Please press any button on the response box to continue.

Vignette 1

Sympathy

Jackie donated $10 to help the victims of a flood. When asked why she donated the money she replied, “I felt enormous sympathy for the victims and I had a really strong desire to help”.

No Sympathy

Jackie donated $10 to help the victims of a flood. When asked why she donated the money she replied, “I had a really strong desire to help”.
Questions:
How morally good was it that Jackie donated money?
1 (morally neutral) – 7 (extremely morally good)

Do you think that Jackie is mainly a good or bad person?
1(mainly a bad person) – 7 (mainly a good person)

Do you think that Jackie has good moral standards?
1(not at all) -7(completely)

Vignette 2

Sympathy

Oliver works with Andrew. It’s lunch time and Andrew is very stressed. He has a lot of work to do and doesn’t have enough time get lunch. Oliver knows that Andrew is stressed and working through lunch, so he buys him a sandwich so he can eat it at his desk. When asked why he bought Andrew a sandwich Oliver replied, “I felt enormous sympathy for him and I had a really strong desire to help”.

No Sympathy

Oliver works with Andrew. It’s lunch time and Andrew is very stressed. He has a lot of work to do and doesn’t have enough time get lunch. Oliver knows that Andrew is stressed and working through lunch, so he buys him a sandwich so he can eat it at his desk. When asked why he bought Andrew a sandwich Oliver replied, “I had a really strong desire to help”.

Questions:
How morally good was it that Oliver bought Andrew a sandwich?

Do you think that Oliver is mainly a good or bad person?
1(mainly a bad person) – 7 (mainly a good person)

Do you think that Oliver has good moral standards?
1(not at all) -7(completely)

Vignette 3

Sympathy

Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked why he visits his friend in the hospital Matthew replied, “I felt enormous sympathy for him and I had a really strong desire to help”.

No Sympathy

Matthew has a friend who was in a car accident and has been recovering in hospital for the last few months. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery.
When asked why he visits his friend in the hospital he replied, “I had a really strong desire to help”

**Questions:**

How morally good is it that Matthew visits his friend in hospital?

Do you think that Matthew is mainly a good or bad person?

1 (mainly a bad person) – 7 (mainly a good person)

Do you think that Matthew has good moral standards?

1 (not at all) - 7 (completely)
Appendix H

Study 5b questioned vignettes

Initial instructions

In this task we are interested in how people evaluate others’ actions. To do this we will present you with several situations and ask you what you think about the actions of the people in each situation.

Each question will be accompanied by a scale. To evaluate an action please press the number on the response box that corresponds to the scale.

Example Question:
Matthias gave Jon $10 for lunch because Jon didn't have enough money to pay for lunch.

How much morally good was it that Matthias gave Jon $10? 1(Morally neutral) – 7(Extremely morally good)

Please read each question carefully. When you are ready to begin press the spacebar to begin.

Instructions 2

In a previous task you were presented with situations and asked to evaluate the actions of people in the situations. In this task you will be presented with the same situations you saw earlier. This time however some of the reasons for why the people behaved in the way they did have CHANGED.

PLEASE READ EACH QUESTION CAREFULLY BEFORE YOU ANSWER. Your response may be the same or different from the first time you read about the situations.

Please press any button on the response box to continue.

Vignette 1

Sympathy
Matthew has a friend who was in a car accident and has been recovering in hospital for the last few weeks. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery. When asked whether he visits his friend because he feels sympathy for him he replied, “I feel enormous sympathy for him and I have a really strong desire to help”.

No Sympathy
Matthew has a friend who was in a car accident and has been recovering in hospital for the last few weeks. Twice a week Matthew goes to visit his friend in the hospital. Matthew’s visits bring great comfort to his friend during his slow and painful recovery.
When asked whether he visits his friend because he feels sympathy for him he replied, “Not really, I just have a really strong desire to help”.

**Questions**

How morally good is it that Matthew visits his friend in hospital each day?
1 (morally neutral) – 7 (extremely morally good)

Do you think that Matthew is mainly a good or bad person?
1 (mainly a bad person) – 7 (mainly a good person)

Do you think that Matthew has good moral standards?
1 (not at all) – 7 (completely)

**Vignette 2**

**Sympathy**

Sharon read that the government has recently cut funding for home care for the elderly. She decides to volunteer to work 1 hour a month helping elderly people with jobs around the house that they have difficulty doing themselves. When asked whether she volunteered because she felt sympathy for the elderly people she replied, “I felt enormous sympathy for them and I had a really strong desire to help”.

**No Sympathy**

Sharon read that the government has recently cut funding for home care for the elderly. She decides to volunteer to work 1 hour a month helping elderly people with jobs around the house that they have difficulty doing themselves. When asked whether she volunteered because she felt sympathy for the elderly people she replied, “Not really, I just had a really strong desire to help”.

**Questions**

How morally good was it that Sharon volunteered to help the elderly people?
1 (morally neutral) – 7 (extremely morally good)

Do you think that Sharon is mainly a good or bad person?
1 (mainly a bad person) – 7 (mainly a good person)

Do you think that Sharon has good moral standards?
1 (not at all) – 7 (completely)

**Vignette 3**

**Sympathy**

Jon just watched a news story about people getting sick because they don’t have access to clean drinking water. Jon decides to donate $5 to an organization that helps provide clean drinking water to those who don’t have access to it. When asked whether he donated the money because he felt sympathy for the people who didn’t have access to clean water he replied, “I felt enormous sympathy for them and I had a really strong desire to help”.

**No Sympathy**

Jon just watched a news story about people getting sick because they don’t have access to clean drinking water. Jon decides to donate $5 to an organization that helps provide clean drinking water to those who don’t have access to it. When asked whether he donated the money because he felt sympathy for the people who didn’t have access to clean water he replied, “Not really, I just had a really strong desire to help”.

142
Questions
How morally good is it that Jon donated $5 to help people without clean drinking water?
1(morally neutral) – 7 (extremely morally good)

Do you think that Jon is mainly a good or bad person?
1(mainly a bad person) – 7(mainly a good person)

Do you think that Jon has good moral standards?
1 (not at all) - 7(completely)
Appendix I

Study 6 materials

Instructions

In this task we are interested in your moral intuitions. You will read a description of someone doing something, and will then be asked three questions about the behaviour of the person. A scale like the one below will accompany each question:

-4 (Extreme blame) – 4(Extreme praise)

Please circle the number on the scale that corresponds to your answer. For example if the question asks “How much praise or blame should this person get for their actions”, and you think they should receive a lot of blame for what they did, you would circle one of the negative numbers on the left hand side of the scale. On the other hand if you think they should be praised for their actions you would circle one of the higher numbers on the right hand side of the scale. And if you think they should be neither praised nor blamed you would circle a number in the middle. Please try to accurately indicate what you think about each person’s behaviour.

Frontal lobe questions

Morally bad

Deliberate: Charlotte deliberately and intentionally punched a student in her class who was rude to her.

Impulsive: Charlotte has frontal lobe damage which results in occasional overwhelming and uncontrollable impulses. Because of one such overwhelming and uncontrollable impulse Charlotte punches a student in her class who is rude to her.

How moral was it that Charlotte punched the student? -4(Extremely morally bad) – 4(Extremely morally good)

How much blame or praise should Charlotte get for punching the student? -4(Extreme blame) to 4(Extreme praise)

How much control do you think Charlotte had over her actions when she punched the student? 0(No control) - 8(Complete control)

Morally good

Deliberate: Jeffrey deliberately and intentionally gave half his savings ($2000) to charity.
Impulsive: Jeffrey has frontal lobe damage, which results in uncontrollable impulsive behaviour. Because of an overwhelming and uncontrollable impulse, Jeffrey impulsively gave half his savings ($2000) to charity.

How moral was it that Jeffrey gave half his savings to charity? -4(Extremely morally bad) – 4(Extremely morally good)

How much blame or praise should Jeffrey get for giving half his savings to charity? -4(Extreme blame) to 4(Extreme praise)

How much control do you think Jeffrey had over his actions when he gave half his savings to charity? 0(No control) - 8(Complete control)

Impulse control questions

Morally bad

Deliberate: Jenna deliberately and intentionally left the restaurant without paying the bill.

Impulsive: Jenna has had impulse control difficulties since she was a child. She occasionally experiences extremely powerful and compelling urges to do things that she can’t control. Because of one such overwhelming and uncontrollable impulse, Jenna impulsively left the restaurant without paying the bill.

How moral was it that Jenna left the restaurant without paying the bill? -4(Extremely morally bad) – 4(Extremely morally good)

How much blame or praise should Jenna get for leaving the restaurant without paying the bill? -4(Extreme blame) to 4(Extreme praise)

How much control do you think Jenna had over her actions when she left the restaurant without paying the bill? 0(No control) - 8(Complete control)

Morally good

Deliberate: While waiting in line at a checkout to pay for her groceries Hilda notices that the lady in front of her doesn’t have enough money to pay for her groceries. Hilda deliberately and intentionally gives the lady $50 to pay for her groceries.

Impulsive: Hilda has had impulse control difficulties since she was a child. She occasionally experiences extremely powerful and compelling urges to do things that she can’t control. While waiting in line at a checkout to pay for her groceries Hilda notices that the lady in front of her doesn’t have enough money to pay for her groceries. Because of a powerful and overwhelming impulse, Hilda impulsively gives the lady $50 to pay for her groceries.
How moral was it that Hilda gave the lady money to pay for her groceries? - 4(Extremely morally bad) – 4(Extremely morally good)

How much blame or praise should Hilda get for giving the lady money to pay for her groceries? -4(Extreme blame) to 4(Extreme praise)

How much control do you think Hilda had over her actions when she gave the lady money to pay for her groceries? 0(No control) - 8(Complete control)
Appendix J

Replication of Asymmetry Effect (Pizarro, Uhlman, & Salovey, 2003)

Method

Participants

94 (f = 44, m = 50) students (M<sub>age</sub> = 22.2 SD = 4.14) studying first or second year psychology at the University of Otago participated in this study in exchange for partial course credit.

Procedure and Materials

The questions relevant to this study were part of a larger study looking at participants’ moral judgements. This larger study was combined with another unrelated study and presented to participants as a pencil and paper questionnaire. Participants completed the study in groups of 3 to 20 participants. Participants gave informed consent before they began the study.

Participants responded to two vignettes taken from Pizarro et al. (2003). In the first vignette an actor performed a morally good action, in the second an actor performed a morally bad action. The order of the vignettes was the same for all participants. For each vignette the actor was described as having performed the morally good or bad action either deliberately and intentionally or uncontrollably and impulsively (see Table A.1). If the morally good action was described as having been performed deliberately then the morally bad action would be described as having been performed uncontrollably and vice versa. Which mental state was associated with the morally good or bad action was manipulated between participants.
Table A.1. *Wording of vignettes for each condition*

<table>
<thead>
<tr>
<th>Mental state (between)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome</strong> (within)</td>
</tr>
<tr>
<td>Morally Good</td>
</tr>
<tr>
<td>Morally Bad</td>
</tr>
</tbody>
</table>

Participants were then asked three questions and asked to respond using a 9 point scale. *How moral or immoral is X's behaviour?* -4 (Extremely immoral) to 4 (Extremely moral). *How much praise or blame should X get for his actions?* -4 (Extreme blame) to 4 (Extreme praise). The third question and scale used by Pizarro et al. (2003) in experiment 1 was omitted for brevity. And a third question asking about responsibility was added but is not included in the following analysis. *How responsible is X for his actions?* -4 (Not responsible at all) to 4 (Completely responsible).

**Results**

Following Pizarro et al. we created a moral judgement score by calculating the absolute value for participant’s responses. The moral judgement score was then calculated for each vignette by taking the average of the immoral-moral and blame-praise questions. *T* tests found no significant difference between deliberate morally good acts (*M* = 2.92, *SD* = 0.90) compared to uncontrollable morally good acts impulsive positive (*M* = 2.83, *SD* = 0.99), *t*(93) = 0.46, 95% CI [-0.30, 0.48]. Deliberate morally bad acts (*M* = 3.42, *SD* = 0.81) judged to be significantly morally worse than impulsive morally bad actions (*M* = 2.89, *SD* = 0.81), *t*(93) = 3.14, 95% CI [0.20, 0.86] (Figure A.1)
Discussion

Consistent with the original findings of Pizarro, Uhlman, & Salovey (2003) impulsive morally wrong bad actions were judged to be less bad than than deliberate morally bad actions. Furthermore consistent with Pizarro et al.'s findings there was no significant difference between deliberate and impulsive morally good actions.
Appendix K

Study 7 materials

Instructions
Next you will be presented with the same situations again. This time however some of the reasons for why the people behaved in the way they did have changed. Please read each question carefully before you answer. Your response may be the same or different from the first time you read about the situations. Please press the spacebar to continue.

Impulse

V1. Charlotte has frontal lobe damage which results in occasional overwhelming and uncontrollable impulses. Because of one such overwhelming and uncontrollable impulse Charlotte punches a student in her class who is rude to her.

V2. Jeffrey has frontal lobe damage, which results in occasional overwhelming and uncontrollable impulses. Because of one such overwhelming and uncontrollable impulse Jeffrey donates $200 to help the victims of a flood.

V3. Robert has frontal lobe damage which results in occasional overwhelming and uncontrollable impulses. Because of one such overwhelming and uncontrollable impulse Robert calls up his local member of parliament and threatens to kill him.

V4. Hilda has frontal lobe damage, which occasionally results in overwhelming and uncontrollable impulses. While waiting in line at a checkout to pay for her groceries Hilda notices that the lady in front of her doesn't have enough money to pay for her groceries. Because of one such overwhelming and uncontrollable impulse Hilda gives the lady $50 to pay for her groceries.

V5. Because of his overwhelming and uncontrollable anger, Ben impulsively smashed the window of the car parked in front of him because it was parked too close to his car.

V6. Melanie is driving her car when a man suddenly steps into the road. Melanie hits the man with her car and he falls face down on to the road. Because of her overwhelming and uncontrollable fear of going to jail Melanie impulsively keeps driving and doesn't check on the man.

Deliberate

V1. Charlotte deliberately and intentionally punches a student in her class who is rude to her.

V2. Jeffrey deliberately and intentionally donates $200 to help the victims of a flood.

V3. Robert deliberately and intentionally calls up his local member of parliament and threatens to kill him.
V4. While waiting in line at the checkout to pay for her groceries Hilda notices that the lady in front of her doesn't have enough money to pay for her groceries. Hilda deliberately and intentionally gives the lady $50 to pay for her groceries.

V5. Ben deliberately and intentionally smashed the window of the car parked in front of him because it was parked too close to his car.

V6. Melanie is driving her car when a man suddenly steps into the road. Melanie hits the man with her car and he falls face down on to the road. Because she doesn't want to go to jail Melanie deliberately and intentionally decides to keep driving and not check on the man.

Questions after each vignette were of the form

V1. How much praise or blame should X get for Y?
-6(Extreme blame) − 0(Neither praise nor blame) − 6(Extreme praise)

How morally good or bad was it that X did Y?
-6(Extremely morally bad) − 0(Neither good nor bad) − 6(Extremely morally good)

How positively or negatively should X be evaluated for Y?
-6(Extremely negatively) − 0(Neither positively nor negatively) − 6(Extremely positively)

To what extent did X really want to Y?
-3(Really did not want to do it) − 0(Neither) − 3(Really wanted to do it)

To what extent do you think X wanted to have an impulse to Y?
-3(Really did not want to have the impulse) − 0(Neither) − 3(Really wanted to have the impulse)
Appendix L

Study 7 supplementary data analysis

Moral goodness score interactions

The four way interaction, control manipulation × morality × vignette × order was not significant $F(1,53) = 0.568$, $p = .454$, $\eta^2_p = .011$. None of the three way interactions were significant, control manipulation × morality × vignette $F(1,53) = 1.322$, $p = .255$, $\eta^2_p = .024$, morality × vignette × order $F(1,53) = 2.646$, $p = .110$, $\eta^2_p = .048$, control manipulation × morality × order $F(1,53) = 1.842$, $p = .18$, $\eta^2_p = .034$, control manipulation × morality × order $F(1,53) = 1.604$, $p = .211$, $\eta^2_p = .029$. The interaction between vignette and control manipulation, $F(1,53) = 4.261$, $p = .44$, $\eta^2_p = .074$, control manipulation and morality, $F(1,53) = 3.380$, $p = .072$, $\eta^2_p = .060$, order and morality, $F(1,53) = 0.076$, $p = .783$, $\eta^2_p = .001$, and vignette and order, $F(1,53) = 1.145$, $p = .289$, $\eta^2_p = .021$, were not significant. There were significant two way interactions between vignette and morality, $F(1,53) = 25.382$, $p < 0.000$, $\eta^2_p = .32$, and order and control manipulation, $F(1,53) = 6.331$, $p = .015$, $\eta^2_p = .015$.

Meta desire analysis

The two questions assessing meta-desires were averaged. A 2(morality: morally good, morally bad; within-subject) x 2(control manipulation: deliberate, impulsive; within-subject) x 2(vignette; within-subjects) x 2(order; between-subjects) ANOVA was conducted.

- morality x control x order x vignette $F(1,66) = .098$, $p = .11$, $\eta^2_p = .74$
- control x vignette x order $F(1,66) = 3.42$, $p = .07$, $\eta^2_p = .05$
- morality x vignette x order $F(1,66) = .99$, $p = .32$, $\eta^2_p = .02$
- morality x control x vignette $F(1,66) = 10.13$, $p < .01$, $\eta^2_p = .13$
- morality x control x order $F(1,66) = 4.34$, $p = .04$, $\eta^2_p = .06$
- morality x order $F(1,66) = .66$, $p = .42$, $\eta^2_p = .01$
- morality x control $F(1,66) = 1.39$, $p = .24$, $\eta^2_p = .02$
- morality x vignette $F(1,66) = .65$, $p = .42$, $\eta^2_p = .01$
- control x vignette $F(1,66) = 2.72$, $p = .10$, $\eta^2_p = .04$
- control x order $F(1,66) = 2.63$, $p = .11$, $\eta^2_p = .04$
- vignette x order $F(1,66) = 2.08$, $p = .15$, $\eta^2_p = .03$
<table>
<thead>
<tr>
<th>Category</th>
<th>F(1,66)</th>
<th>p</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>control</td>
<td>152.35</td>
<td>&lt; .01</td>
<td>.70</td>
</tr>
<tr>
<td>moral</td>
<td>48.52</td>
<td>&lt; .01</td>
<td>.42</td>
</tr>
<tr>
<td>order</td>
<td>.274</td>
<td>.60</td>
<td>.00</td>
</tr>
<tr>
<td>vignette</td>
<td>2.96</td>
<td>.09</td>
<td>.04</td>
</tr>
<tr>
<td>Morally good</td>
<td>t(67) = 13.60</td>
<td>&lt; .01</td>
<td>d = 1.96</td>
</tr>
<tr>
<td>Morally bad</td>
<td>t(67) = 8.57</td>
<td>&lt; .01</td>
<td>d = 1.46</td>
</tr>
</tbody>
</table>
Appendix M

Study 8 materials

Instructions

In this task we are interested in your moral intuitions. You will read a description of someone doing something, and will then be asked three questions about the behaviour of the person. A scale like the one below will accompany each question:

-4 (Extreme blame) – 4(Extreme praise)

Please circle the number on the scale that corresponds to your answer. For example if the question asks “How much praise or blame should this person get for their actions”, and you think they should receive a lot of blame for what they did, you would circle one of the negative numbers on the left hand side of the scale. On the other hand if you think they should be praised for their actions you would circle one of the higher numbers on the right hand side of the scale. And if you think they should be neither praised nor blamed you would circle a number in the middle. Please try to accurately indicate what you think about each person’s behaviour.

V1

Anger

Because of her overwhelming and uncontrollable anger at discovering the company she worked for was dumping chemicals illegally into a local stream, Betty impulsively reported the company to the police.

Deliberate

Betty discovered that the company she worked for was dumping chemicals illegally into a local stream. Betty deliberately and intentionally reported the company to the police.

How praiseworthy was it that Betty called the police?
1(Not at all praiseworthy) – 7(Extremely praiseworthy)

How morally good was it that Betty called the police?
1(Not at all morally good) – 7(Extremely morally good)

How positively should Betty be evaluated for calling the police?
1(Not at all positively) – 7 (Extremely positively)

V2

Anger

Jack watches as several people ignore a homeless man sheltering from the cold. Because of his overwhelming and uncontrollable anger at the people ignoring the homeless man, Jack impulsively gave the homeless man his only jacket even though it was freezing outside.

Deliberate
Jack watches as several people ignore a homeless man sheltering from the cold. Jack deliberately and intentionally gave the homeless man his only jacket even though it was freezing outside.

How praiseworthy was it that Jack gave the homeless man his jacket?  
1(Not at all praiseworthy) – 7(Extremely praiseworthy)

How morally good was it that Jack gave the homeless man his jacket?  
1(Not at all morally good) – 7(Extremely morally good)

How positively should Jack be evaluated for giving the homeless man his jacket?  
1(Not at all positively) – 7 (Extremely positively)

V3  
Anger  
Lucy just heard that the government has cut funding to the local hospital. Because of her overwhelming and uncontrollable anger at the government for cutting funding Lucy impulsively donates $300 to the hospital.  
Deliberate  
Lucy just heard that the government has cut funding to the local hospital. Lucy deliberately and intentionally donates $300 to the hospital.

How praiseworthy was it that Lucy donated $300 to the hospital?  
1(Not at all praiseworthy) – 7(Extremely praiseworthy)

How morally good was it that Lucy donated $300 to the hospital?  
1(Not at all morally good) – 7(Extremely morally good)

How positively should Lucy be evaluated for donating $300 to the hospital?  
1(Not at all positively) – 7 (Extremely positively)
Appendix N

Studies on disgust and fear

The following study is an examination of the influence of uncontrollable anger, fear and disgust on blame judgements. It is hypothesized that uncontrollable anger reduces control while uncontrollable fear and disgust do not, and so blame will be discounted for uncontrollable anger but not fear or disgust.

Method

Participants

See study 7 for anger and fear participants and study 6 for disgust participants.

Procedure and Materials

The fear and anger questions were run at the same time as Study 7 and the vignettes were presented after the vignettes for Study 7. For design see Study 7. The disgust question was asked at the same time as Study 6, see Table A.2 for vignettes.

Table A.2. Vignettes used to examine the influence of anger, fear, and disgust on blame judgements.

<table>
<thead>
<tr>
<th>Morality</th>
<th>Deliberate</th>
<th>Uncontrollable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>Betty discovered that the company she worked for was dumping chemicals illegally into a local stream. Betty deliberately and intentionally reported the company to the police. Sam made a mistake at work that will cost the company he works for hundreds of thousands of dollars. Because he doesn't want to lose his job, Sam deliberately and intentionally falsifies several documents to make it look like one of his co-workers made the mistake and not him.</td>
<td>Because of his overwhelming and uncontrollable anger, Ben impulsively smashed the window of the car parked in front of him because it was parked too close to his car. Sam made a mistake at work that will cost the company he works for hundreds of thousands of dollars. Because he doesn't want to lose his job, Sam deliberately falsifies several documents to make it look like one of his co-workers made the mistake and not him.</td>
</tr>
</tbody>
</table>
Disgust

Sarah deliberately and intentionally spat on the protester because she fundamentally disagreed with what he was saying. Because of her overwhelming and uncontrollable disgust, Sarah impulsively spat on the protester because she fundamentally disagreed with what he was saying.

Results

Fear

2 (emotion: fear, anger; within) x 2 (control manipulation: deliberate, impulsive) x 2 (order: between). The three way interaction, emotion x control manipulation x order was not significant $F(1,60) = .880$, $p = .352$, $\eta^2_p = .0014$. The interactions between order and the other factors were not significant, order x control manipulation $F(1,60) = 1.654$, $p = .203$, $\eta^2_p = .027$, order x emotion $F(1,60) = 1.054$, $p = .309$, $\eta^2_p = .017$. The interaction between control manipulation and emotion was significant, $F(1,60) = 14.277$, $p < .000$, $\eta^2_p = .19$. Angry actions ($M = 4.53$, $SD = 1.28$) received less blame than did fearful actions ($M = 5.15$, $SD = 1.13$), with a significant main effect for emotion type, $F(1,60) = 19.242$, $p < .000$, $\eta^2_p = .243^{40}$. Deliberate actions ($M = 5.08$, $SD = 1.13$) received significantly more blame than did uncontrollable actions ($M = 4.60$, $SD = 1.28$), $F(1,60) = 34.014$, $p < .000$, $\eta^2_p = .362$. The order main effect was also significant, $F(1,60) = 5.550$, $p = .022$, $\eta^2_p = .085$ with blame greater for order1 ($M = 5.04$, $SD = 1.13$) than order2 ($M = 4.62$, $SD = 1.25$).

For the vignette pair where uncontrolled anger was the uncontrolled emotion the deliberate vignette ($M = 5.02$, $SD = 1.06$) received significantly higher blame than did the impulsive anger vignette ($M = 4.11$, $SD = 1.49$), $t(62) = 5.310$ 97.5% CI [0.52, 1.30], $d = 0.70$. For the vignette pair that used fear as the uncontrollable emotion, the deliberate vignette ($M = 5.14$, $SD = 1.05$) did not receive significantly higher sanction scores than did the impulsive fear vignette ($M = 5.21$, $SD = 1.12$), $t(65) = 0.672$ 97.5% CI [-0.29, 0.16], $d = 0.06$ (See Figure A.2).

Disgust

A $t$ test was conducted to compare control manipulations for a single vignette pair. There was no significant difference between the vignette that described the agent acting deliberately ($M = 2.71$, $SD = 1.12$) and acting because of uncontrollable disgust ($M = 2.84$, $SD = 1.01$), $t(117) = 0.634$, 95% CI [-0.51, 0.26]

Discussion

The results are inconsistent with the asymmetry hypothesis. Blame was discounted for uncontrollable anger but not uncontrollable fear or disgust.

Power to detect an effect of control on moral judgements for the fear question was adequate to detect medium effect ($d = .5$) and larger ($power = .99$), power was not adequate to detect effect sizes 0.3 and smaller. We can estimate the effect size of

---

40 This is because of the interaction where blame was discounted for uncontrolled angry actions
uncontrollable anger from study 3 as .77, and from Pizarro et al. Study 1 \( d = .80 \), study 2 \( d = 1.26 \), study 3 \( d = 1.02 \). So we should have adequate power to detect an effect that is close to the anger effect size. For the disgust question power to detect an effect size of \( d = .3 \) and larger was .94, so again we can't appeal to lack of power.

These two questions provide further evidence that uncontrollable anger reduces control while other uncontrollable emotions, disgust, fear, and sympathy do not.

![Figure A.2](image_url)

*Figure A.2.* Mean moral judgements scores (+SE) deliberate and uncontrollable actions. The uncontrollable emotion is either anger, fear or disgust. Vignettes were described as deliberate or uncontrolled. Each uncontrollable emotion is graphed next to its deliberate counterpart.
Pizarro et al. (2003) hypothesized that the asymmetry effect is due to assumptions observers make about people's second-order desires (or meta-desires). According to the authors, second-order desires are an individual’s acceptance or rejection of a desire or impulse. For example for a drug addict “there is a second-order desire (not to be an addict) and a first-order desire (a compulsion to continue to take drugs)” (Pizarro et al, p267).

To assess blame for an uncontrollable action the authors argue it is reasonable to consider whether an individual endorses the actions they have been compelled to commit. The authors argue that the naïve observer has expectations about second-order desires. For an individual impulsively committing a morally bad action an observer has the expectation that the individual has a second-order desire not to commit the morally bad action. Because the second-order desire is inconsistent with the first-order impulse the amount of blame the individual receives is reduced because an observer infers that the individual really didn't want to commit the morally bad action, they only did it because of an impulse. On the other hand when an individual impulsively commits a morally good action observers have the expectation that the individual also has a second-order desire to perform the morally good action. And so the first-order impulse to perform the morally good action and the inferred second-order desire to also perform the morally good action are the same, and so discounting for a loss of control is not needed. Pizarro et al. formalise this line of reasoning as the meta-desire hypothesis: “The asymmetry can be explained by appealing to the assumptions held by naïve judges about the agent's second-order desires. Specifically, we hypothesized that naïve judges assume that impulsive negative acts are accompanied by conflicting (positive) second-order desires, but that impulsive positive acts are accompanied by consistent (positive) second-order desires. The belief that individuals consistently possess positive second-order desires may be the driving force behind the hypothesized asymmetry” (p.267).

Pizarro et al. ran two follow up studies to examine the role of second-order desires in the asymmetry. In study 2 the authors ran a similar study to study 1 which had participants evaluate vignettes that had agents either impulsively or deliberately
performing morally good or bad actions. Study 2 included a third group where agents acting because of an impulse, and participants were additionally told that the agent would rather not have had the impulse (the meta-desire condition). It was hypothesised that when described as having second-order desires that are inconsistent with their first-order impulses participants would discount blame and praise for both morally good and bad actions. The authors reported that the results were consistent with this hypothesis.

In study three Pizarro et al. again looked at the role of second-order desires. This time instead of informing the participant about the agent's second-order desires, participants were asked what they thought the agent’s second-order desires were. The authors then ran a mediation analysis and predicted that participants perceived second-order desires would mediate the asymmetry effect. The results were consistent with their predictions and they conclude that the asymmetry is mediated by second-order desires.

According to Pizarro et al. a second-order desire “can be defined as an individual's higher-order acceptance or rejection of a desire or an impulse”. This is consistent with the questions the authors ask participants in Study 3: “This person would rather not have an impulse to ..., to what extent do you think this person wanted to have an impulse to ..., and the manipulation in their second study, “But wished that he did not possess such an impulse”.

However the way the authors talk about second-order desires is not entirely consistent, especially when they talk about how the meta-desire account of the asymmetry works. “In the case of a drug addict, for instance, there is a second-order desire (not to be an addict) and a first-order desire (a compulsion to continue to take drugs)”41, 42.” Here there is a second-order desire “not to be an addict” which expresses their position towards being addicted, and not directly acceptance or rejection of an impulse. And later “However, in the case of a positive act, individuals likely assume that the second-order desire is consistent with the first-order impulse, and thus see no need to discount moral praise.” If a second-order desire is the acceptance or rejection of

41 The second-order desire described here “not to be an addict” isn't an expression of acceptance or rejection of a specific impulse but desire that implies rejection of an impulse but isn't itself a rejection of a particular impulse, but maybe this isn't a problem.
42 Actually it sounds like the desire they are talking about is a general sort of desire and isn't specific to an event. And the second-order desire is also general but isn't specific to an evaluation of the first-order desire but is inconsistent with it
a first-order impulse then it doesn't make sense to talk about the first-order impulse and the second-order desire being *consistent*. It makes sense to talk about first and second-order desires being consistent if second-order desires are desires that expresses an evaluation of, for example, a course of action that could be consistent with a first-order impulse. The third question of Study 3 which seeks to probe the participants assumptions about an agents second-order desires, asks “To what extent did this person really want to do what he did?”, again this isn't a question about acceptance or rejection of an impulse. Acceptance or rejection of the impulse is implied by their answer, but this question seems to be getting at an agents 'true' desires, “*Inform them about the true desire of the agent*” (p. 267).

Pizarro et al. appear to be talking about second-order desires in two different ways, as acceptance or rejection of an impulse, and an agents true desire. Apart from the fact that an account of meta-desires needs to be clear if we are going to make predictions it has bearing on interpretation of our results. The question is can the meta-desire hypothesis account for the results of Studies 6-8.

The meta-desire hypothesis argues that praise is not discounted in the sympathy condition because observers assume that the agent's second order desires are consistent with performing a morally good action. If we consider meta-desires as acceptance or rejection of an impulse, then it is an open question whether participants think that the agent would endorse an impulse they had due to brain damage. Perhaps they would not accept the impulse because impulses due to brain damage are not something a person wants to have. But maybe participants think that the agent would accept the impulse because it leads to a morally good action.

If meta-desires on the other hand are the true desires of an agent, then there does not appear to be any reason to expect observers to make different assumptions about an agents second-order desires, regardless of whether the impulse is from emotion of brain damage, because the assumptions observers make, according to the meta-desire hypothesis, are dependent on the morality of the action. And so if an observer makes assumptions about second-order desires in a similar fashion for impulses caused by frontal lobe damage as sympathy then the meta-desire hypothesis can not account for the results from Study 6 and 7.

A meta-desire interpretation of Study 8 is similar. Observers may assume that an
agent would not endorse an angry impulse if the observer considers angry impulses in this context to be negative—though anger or moral outrage could be evaluated positively. Or if the observer is trying to work out what the true desire of the agent is then we would expect that they would again assume that the agent's second-order desires are consistent with the first-order angry impulse because the action is morally good.

The account of meta-desires is not clear enough to make a clear prediction as to whether observers are likely to make the same kinds of meta-desire assumptions for morally good actions resulting from uncontrollable sympathy, uncontrollable anger, and frontal lobe damage. This is an empirical question anyway, but we may not have to spend time thinking about this or running studies to work out how people are responding to these second-order desire questions because there is reason to doubt the empirical evidence for the meta-desire interpretation of the asymmetry evidence.

**Empirical evidence**

Evidence for the meta-desire hypothesis is inconclusive. Pizarro et al. conducted two studies in support of the meta-desire hypothesis. The first study was similar to the original asymmetry study with the addition of a meta-desire condition. The meta-desire condition was identical to the impulsive condition except that participant were told that the agent wishes that they didn't have the impulse to perform the action. The design is 2(valence: positive, negative) x 3 (control manipulation: deliberate, impulsive, meta-desire) x 2(vignette). The authors predicted that “if participants were provided with knowledge that an agent who felt compelled to commit a positive act possessed a second order desire not to commit this act, the asymmetry would disappear – moral sanctions would be discounted for positive impulsive acts, as they are for negative impulsive acts” (p. 269).

The authors ran the same ANOVA as they had in study one, and followed up the ANOVA in the same way with a simple effects comparisons between deliberate and impulsive morally good actions and then deliberate and impulsive morally bad actions. However, they don't provide data on comparisons between impulsive and meta-desire conditions which would be the most informative comparison. Instead they run a
separate ANOVA and combine the deliberate and impulsive conditions of the mental-state (called control manipulation in this chapter). The new analysis is 2(morality: morally good, morally bad) x 2(control manipulation: combined impulsive/deliberate, meta-desire). The authors report a significant main effect for the meta-desire condition and report that this indicated that providing information that an agent rejects their own positive impulse resulted in discounting of praise. This analysis is odd because the composite deliberate, impulsive score is difficult to interpret.

The authors conclude that the results are consistent with the idea that meta-desire information leads to discounting of praise and blame. The results do support this hypothesis for the general effect of inconsistent meta-desire information reducing praise and blame. However the lack of an interaction between the composite score, meta-desire, and morally good and bad actions is actually evidence against the meta-desire hypothesis account of the asymmetry.

The composite score for morally good actions combines deliberate and impulsive actions. According to the meta-desire hypothesis observers assume for impulsive morally good actions that the agent has a second-order desire that is consistent with the first-order impulse. So the composite score is a composite of deliberate actions (where the observer assumes that the agent wants to perform the action because that is what a deliberate action means) and impulsive actions (where the observer assumes that the agent wanted to perform the action anyway: meta-desire hypothesis). When we compare this composite score to the impulsive meta-desire condition where participants are told that the agent doesn't want to perform the morally good action we should expect a large discounting of praise. Now consider the composite scores for morally bad actions. In the deliberate morally bad condition, as in the deliberate morally good condition the agent wants to perform the action. However for the impulsive morally bad condition an observer assumes that the agent doesn't really want to perform the morally bad action. According to the meta-desire hypothesis this condition is the same as the meta-desire condition for morally bad actions. In both the impulsive morally bad condition and the meta-desire morally bad conditions agents are performing morally bad actions but they have second-order desires not to commit them. And so when we compare a composite of deliberate and impulsive morally bad actions with the meta-desire condition for morally bad actions we should expect a smaller
discounting of blame than for morally good actions because the composite score for morally bad actions is already partially discounted because of the impulsive morally bad actions that make up part of the score (see Figure A.2).

![Figure A.2](image)

**Figure A.2.** The relationships predicted by the meta-desire hypothesis between composite moral sanction scores and meta-desire more sanction scores.

And so if the results of this study were consistent with the meta-desire hypothesis they should find a significant two way interaction. The authors did find that in general the presence of inconsistent meta-desire information reduces praise and blame; which shows that meta-desire information reduces praise and blame when it is explicitly given to participants. However the fact that there was not an interaction between the composite scores and the impulsive meta-desire condition is evidence against the meta-desire hypothesis.

Study 3 took a different approach and asked participants whether the agent wanted to have an impulse to perform morally good and bad actions. Study 3 was identical to study 1 with the addition of questions about second-order desires. Participants answered three questions asking them how much they thought the agent wanted to have the impulse to perform an action: “This person would rather not have an impulse to [perform the indicated behavior], To what extent do you think this person wanted to have an impulse to [perform the indicated behavior], To what extent did this person really want to do what he did?”. Participants responded using a 9 point scale: 1
(really did not want to possess), 5 neither, 9 (really did want to have). Responses to the three questions were averaged for each vignette. The authors report that perceived second-order desires were consistent with first-order impulses for morally good actions ($M = 6.8$), but that second-order desires were conflicting with first-order impulses for morally bad actions ($M = 4.8$). I agree that this seems to be the case for morally good actions but the mean judged second-order desires for morally bad actions was 4.8. On the 9 point scale used that places the average rating right in the middle of the scale, which is marked with neither. To interpret these results as indicating that the participants judged the agents second-order desires as being inconsistent with their first-order impulses for morally bad actions is misleading. A mean rating in the middle of the scale indicates that participants either weren't sure what the agents second-order desires were or that they thought the agent had ambivalent second-order desires.

Another question is how participants were answering these questions if both deliberate morally good and bad actions had mean ratings between 6 and 6.5 on the scale. These scores are on the lower end of the side of the 9 point scale that indicates that the agent wanted to perform the action. With participants not very confident in affirming that agents wanted to deliberately perform actions it raises questions about the measure of second-order desires being used. When we are talking about participants making assumptions about agents second-order desires as being either consistent or inconsistent with first-order impulses a rating of slightly above neither doesn't indicate that they are making very strong assumptions at all. It also seems like participants were asked if agents wanted to have an impulse to perform a deliberate action. Asking participants whether someone wanted to have an impulse to perform a deliberate action is an odd question, and odd questions result in odd answers. The results could still be consistent with a modification to the meta-desire hypothesis: For morally good actions there is an assumption that people have positive second-order desires, but that for morally bad actions observers are not sure what a person's second-order desires are.43

The authors then ran a mediation analysis which indicated that second-order desires mediated the relationship between control manipulation and morality44. An alternative interpretation is that participants base their meta-desire judgements –

43 **This is kind of interesting because bad behaviours are usually most diagnostic** (Baumeister et al., 2001).
44 **They use the .. technique which has fallen out of favour.**
whether the agent wanted to perform the action or not- on how much control the agent has over their actions. For example in Study 7 the extremity of moral judgements in the impulsive condition was reduced for both morally good and bad actions; meta-desire judgements – unlike in Pizarro et al. – indicated that participants thought that agents had less desire to perform an impulsive than a deliberate action for both morally good and bad actions. This is consistent with the meta-desire hypothesis and accordingly it could be argued that we did not observe an asymmetry in study 7 because participants did not infer that agents had a similar desire to perform both deliberate and impulsive morally good actions as found in Pizarro et al. However perhaps a more compelling interpretation is that in Study 7 participants inferred that agents had less desire to perform impulsive morally good actions than deliberate morally good actions because unlike in Pizarro et al. Study 7 successfully manipulated control between impulsive and deliberate conditions for morally good actions. Meta-desire judgements are then an indication that control was successfully manipulated between conditions rather than a mediator of the interaction between morality and control manipulation.