**From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Philosophy

---

2018

# Empathic Concern, Moral Epistemology, and Moral Progress

Lerner, Adam Jared

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01rj4307242

*U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf*

# EMPATHIC CONCERN, MORAL EPISTEMOLOGY, AND MORAL PROGRESS

**Adam Jared Lerner**

**A DISSERTATION**

**PRESENTED TO THE FACULTY**

**OF PRINCETON UNIVERSITY**

**IN CANDIDACY FOR THE DEGREE**

**OF DOCTOR OF PHILOSOPHY**

**RECOMMENDED FOR ACCEPTANCE**

**BY THE DEPARTMENT OF**

**PHILOSOPHY**

**Advisers: Mark Johnston and Sarah-Jane Leslie**

**April 2018**

© Copyright by Adam Jared Lerner, 2018. All rights reserved.

**Abstract**

In this dissertation, I argue that a specific form of empathy—*empathic concern*—has a crucial role to play in moral inquiry. I argue that experiences of empathic concern serve as a basic source of evidence regarding the strength of our reasons to help and not harm others, and that recognizing this can help us make progress on debates about the strength of our obligations to reduce animal suffering and extreme poverty.

In chapter 1, I draw on work in cognitive science to argue that we would not be doxastically justified in holding certain moral beliefs unless empathic concern provided evidence of our reasons to help and not harm others. Since we are in fact doxastically justified in holding these beliefs, it follows that empathic concern must provide evidence of this sort.

In chapter 2, I provide an account of how the evidence provided by empathic concern can improve moral judgment despite its various limits. I argue that just as visual perception can be a useful source of evidence despite its inherent limits, empathic concern can provide a useful source of evidence despite its inherent limits. I go on to provide an account of how we should go about collecting this evidence. I argue that to collect the evidence provided by empathic concern in the proper way, we would have to put ourselves in the favorable conditions identified at the end of chapter 1 whenever we harm someone or fail to help them.

In chapter 3, I argue that we have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way, we would judge that we're morally required to consume far fewer animal products than most of us do. I argue in chapter 4 that we have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way, we would judge that we're morally required to give much more money to charity

than most of us do. I argue further that we should defer to the judgments we would make under these conditions.

## Acknowledgments

First, thank you to my advisors, Sarah-Jane Leslie and Mark Johnston, for their unwavering support and guidance. I am profoundly grateful to have had them as advisors. They did not merely tolerate, but actively encouraged the interdisciplinary nature of my work. And they always encouraged me to be ambitious, while at the same time making sure I set realistic expectations. Most importantly, they were steadfast in their support of my work even when crippling anxiety and self-doubt left me floundering. This has made all the difference.

Beyond my dissertation, Sarah-Jane has played a formative role shaping my graduate school career. I am grateful to Sarah-Jane not only for advising me for most of my time in graduate school, but for generously inviting me to collaborate with her as a first-year graduate student. I can imagine no better way to learn how to write philosophy than by doing it with Sarah-Jane as a co-author. Thank you, Sarah-Jane.

Crucial turning points in my philosophical thinking coincided with meetings with Mark, and this is no accident. Meetings with Mark were philosophical journeys, leaving me with a host of questions I would need to reckon with if I wanted to adequately defend my views. I owe many of the ideas I consider in my dissertation and much of what I plan to work on after my dissertation to questions posed by Mark. Thank you, Mark.

Thank you to Nat Tabris. It is no exaggeration to call Nat my unofficial advisor. Nat played an instrumental role at every stage of this process. From helping me articulate my own inchoate thoughts, to helping me understand the structure of my own arguments, and to even formatting the written dissertation itself, Nat was always there, unflinching in his willingness to help. He is a model philosopher, extraordinarily patient, kind, and charitable. He always takes his

interlocutor's ideas seriously, doing his best to understand the motivations for them, and providing constructive feedback in a careful, collaborative manner. Whatever positive qualities this dissertation possesses, it would not possess them to the same degree—or at all—if it were not for Nat. Thank you, Nat.

Thank you to my readers, Johann Frick and Sarah McGrath, both for their detailed feedback on my dissertation which led to important improvements, as well as for regularly going beyond the call of duty during my graduate school career. I was never officially advised by Johann or Sarah, but given how generous they were with their time and guidance, one could be mistaken in thinking otherwise.

I was told as a prospective graduate student that I would learn the most from my fellow graduate students. This has no doubt been true. In addition to Nat, I have learned an enormous amount from Eric Hubble, Robbie Hirsch, and Yoaav Isaacs. And I would have never made it through logic if it were not for Yoaav's patience and generosity.

This dissertation has been improved by conversations, questions, and encouragement from many people, including Merrick Anderson, Dominic Burkart, Dan Burnston, Daniel Drucker, Adam Elga, Ian Fishback, Sam Fullhart, Adam Gibbons, Sherif Girgis, Eleanor Gordon-Smith, Jesse Holloway, Alex Ioakimidis, Tyler John, Jim Joyce, Brett Karlan, Tom Kelly, Victor Kumar, Aaron Kurosu, Daniel Kranzelbinder, Kyle Landrum, Harvey Lederman, Yimao Liu, Alex Meehan, Judith Mildner, Brendan Mooney, Joseph Moore, Dylan Murray, Alexander Nehamas, Tony Phan, Sam Preston, Chris Register, Domenica Romagni, Gideon Rosen, Oliver Sensen, Joy Shim, Dave Shoemaker, Michael Smith, David Schroeren, Oliver Sensen, Michael Smith, Julius Solatorio, Laura Soter, Alvaro Sotil, Pasha Temkin, Mark Thornton, Danny Underwood, Chad Van Schoelandt, Aarthy Vaidyanathan, Elise Woodard,

Frank Wu, Yuqing Yang, Erik Zhang, and Zidong Zhao. I am particularly grateful to Zoë Johnson-King for pointing out a non-obvious but important tension in my master argument that required a significant amount of effort to address. I received much of this feedback during presentations I gave to the Princeton-Michigan Meta-Normativity Graduate Workshop on August 27th, 2017, to a meeting of the philosophy dissertation seminar at Princeton on October 13th, 2017, to a meeting of the Princeton Social Neuroscience Lab on November 20th, 2017, and to the Princeton Philosophy Department on November 20th, 2017.

I am especially grateful to Steve Stich for providing numerous insightful questions and for giving me the opportunity to give an extended presentation of my work in his graduate seminar on moral psychology in October 17th, 2017. I am also especially grateful to Peter Railton, who took time out of his busy schedule to have a lengthy meeting with me when I visited Michigan in August 2017, and who graciously entertained many additional questions I had following a talk he gave at Princeton in November. Thank you to Shaun Nichols whose own work provided an important precedent for my own, and for a very helpful conversation over coffee. Thank you very much to Chris Freiman for very helpful discussion and collaboration, and for his generous mentorship over the years. Thank you to Peter Singer whose own arguments provide much of the inspiration for the second half of this dissertation. Thank you to Paul Bloom for several extended conversations. Thank you to Heidi Maibom for exposing me to much of the philosophically significant empirical work on empathy through her own writing on empathy. Thank you to all of those cognitive scientists whose work made this dissertation possible. I am especially indebted to Daniel Batson, who kindly answered some questions for me over email, and whose corpus of work on perspective-taking, empathic concern, and prosocial behavior first demonstrated to me the power of empathy. I am also especially indebted to James Blair, whose

work on the role of empathic dysfunction in psychopathy inspired my own views on the topic, and Abigail Marsh, whose work on fear recognition, psychopathy, and prosocial behavior informs many of the empirical claims I make in my dissertation.

Thank you to the Graduate School for awarding me a Jacobus Fellowship and a Dissertation Completion Fellowship. Thank you to the University Center for Human Values for providing me a Graduate Prize Fellowship. Lastly, thank you to the Sperry Fund for providing me with a Beinecke Scholarship, which enabled me to take risks I would have not been able to, the first of which was applying to Princeton.

Thank you to my parents, Amy and Steven, for the enormous support they've provided me getting me to this stage in my life. I cannot summarize their support here without seeming to trivialize it. Thank you to Chris, Sheila, Ed, and especially Diana Cimino for their support throughout the dissertation process. Lastly, thank you to my friends, especially those who work tirelessly on behalf of factory-farmed animals. Your exceptional capacity for empathy inspired much of what went into this dissertation.

I am fortunate to have had dozens of additional teachers and friends who have indirectly contributed in myriad ways to the causal chain that culminated in this dissertation. To those who contributed in these indirect ways, I thank you. And to others who contributed to this project in more direct ways who I have accidentally omitted, I apologize and thank you.

## Table of Contents

**Abstract.....iii**

**Acknowledgments.....v**

**Chapter 1: Putting Empathic Concern in Its Place.....1**

**Chapter 2: How to Collect Empathic Evidence.....49**

**Chapter 3: Empathic Concern and Eating Animals.....77**

**Chapter 4: Empathic Concern and the Demands of Morality.....104**

**References.....135**

## Chapter 1

## Putting Empathic Concern in Its Place

"When the suffering of another creature causes you to feel pain, do not submit to the initial desire to flee from the suffering one..."

-Leo Tolstoy

It's 3:00 am and you can't sleep. So you're watching reruns of your favorite sitcom from yesteryear. The show cuts to commercial. The screen fades to black, and you hear a voice emanating from the tinny speakers of your television set. It's Sarah McLachlan. She sings: "In the arms of the angel, far away from here…." If you lived in North America around the turn of the century, you know what's about to happen. If you don't turn the channel, you'll be bombarded by slow-motion video clips zooming in on the faces of neglected and abused dogs and cats. They're going to look at you through the television screen with the widest of eyes and the most sullen of faces, imploring you to help them. You're going to feel an intense mix of emotions well up inside of you: sympathy, compassion, heartbreak, warmth, tenderness. And for all you know, these emotions may well lead you to think you ought to whip out your credit card and phone in a donation to the SPCA.

Now suppose you can freeze time in the second before the kittens and puppies appear on your screen. Suppose further that in this moment your only goal is to figure out whether and how much money you should donate to the SPCA this year. The answer may be nothing at all, or it may be quite a bit. You don't know. But you want to know. So now you have to ask yourself: do

you turn the channel, or do you expose yourself to Sarah McLachlan and her battery of cats and dogs? Would allowing yourself to feel *empathic concern* for these cats and dogs help or hinder you in your quest to discover how much you should help?

This question is a specific instance of a more general question that has long fascinated philosophers: what is the proper role of emotions in moral inquiry? If we want to know how we should lead our lives, will emotions help us gain this knowledge? In the first half of my dissertation, I explore a restricted version of this question: what legitimate role, if any, does empathic concern play in helping us understand what's right and what's wrong? In the second half of the dissertation, I argue that appreciating the role of empathic concern can help us make progress on two seemingly intractable debates about the scope and strength of our moral obligations to reduce suffering. In particular, it helps us resolve debates about the moral significance of animal suffering and the moral obligations of affluent individuals to give to charity.

"Empathic concern" refers to the kinds of emotional reactions we would expect someone to experience when thinking about the misfortune of someone they value for their own sake. The specific form that empathic concern takes in any given situation varies depending on the nature of the misfortune that triggered it. When the misfortune has already occurred, empathic concern may closely resemble sadness. When the misfortune is in the future, empathic concern may more closely resemble fear. But whatever particular form it takes, empathic concern is often infused with pleasant feelings of "tenderness," "warmth," and "compassion" (Lopez-Perez et al. 2014). Among the forms empathic concern can take, psychologist Daniel Batson lists "sympathy, compassion, softheartedness, sorrow, sadness, upset, distress, concern, and grief" (2011, p. 12).

Why should we care what role empathic concern has to play in helping us understand what's right and what's wrong? We should care because empathic concern exerts a powerful influence over our moral judgments, and we need to know whether we should welcome this influence. We need to know whether we should seek it out or go out of our way to resist it. With the proliferation of media tailor-made to induce empathic concern, we now face these kinds of decisions on a daily basis. Should we watch the SPCA ad or change the channel? Should we watch the video of the police officer brutally beating the innocent protester? Should we follow the link to the factory farm footage that our vegan friend posted on the internet? Should we strap on a virtual reality headset and transport ourselves to a refugee camp? How we answer these questions may have a significant influence on how our moral attitudes evolve in the future. So we need to make sure we answer these questions correctly. But we cannot answer these questions correctly until we know what legitimate role empathic concern has to play in moral inquiry.

The epistemic significance of empathic concern is not limited to answering urgent questions about how we should regulate the influence of empathic concern on the future evolution of our moral beliefs. It's also important for evaluating the epistemic status of the moral beliefs we already have. As I will argue in later sections of this chapter, empathic concern has an important role to play in moral learning, and the patterns of empathic concern we experience over the course of our moral development have a significant role to play in fixing the content of our moral beliefs. For that reason, we need to know whether certain patterns of empathic concern are preferable to others. It may turn out that certain patterns of empathic concern are preferable to the actual patterns that influenced our moral development. If so, we should adopt the moral judgments we have sufficient reason to believe we would have made had we instead experienced these preferable patterns of empathic concern.

Let me put my cards on the table. I believe empathic concern has a significant and underappreciated role to play in moral inquiry. I also believe that we feel far too little empathic concern and feel it far too infrequently, and this seriously compromises our ability to know how we should live our lives. In today's globalized economy, we do not regularly expose ourselves to the kinds of stimuli that would lead us to feel empathic concern for those we regularly harm and fail to help. We do not typically feel empathic concern for the child whose life we could save by donating money to charity rather than buying a fancy car. This child is not present when we're signing the papers in the lobby of the car dealership. They are not sitting in the passenger seat, coughing and wheezing as we drive off the lot. Neither do we feel empathic concern for the animals who have had to suffer to produce the animal products that constitute so much of our diets. Our pork chop does not scream when we tuck into it. Our milk does not sob as we sip it.

I argue that because we constantly fail to feel adequate empathic concern for these suffering individuals, we are constantly missing key evidence regarding the strength of our reasons to help and not harm them. The fact that we are missing this evidence profoundly undermines our ability to discover how we should treat them. It leads us to radically underestimate the strength of our obligations to give money to charity and to avoid supporting industrial animal agriculture. Most of us should be eating far fewer animal products and giving far more to charity than we currently do.

The arguments for these first-order moral claims will come in later chapters. In this first chapter, I lay the epistemic groundwork for those arguments. I argue that empathic concern has an ineliminable role to play in moral inquiry: it serves as a basic source of evidence regarding the strength of our reasons to help and not harm others. As we feel more empathic concern for someone, we come to have evidence that our reasons to help and not harm them are stronger.

In section 1, I outline the master argument that empathic concern *at least sometimes* provides evidence regarding the strength of our reasons to help and not harm others. This argument depends crucially on what I call the *Empathic Foundations Hypothesis*. According to the Empathic Foundations Hypothesis, we only believe we have strong reasons to help and not harm others because we, or our testimonial informants, have in the past treated empathic concern as providing us with evidence regarding the strength of our reasons to help and not harm others. The next three sections provide an extended argument for the Empathic Foundations Hypothesis. In section 2, I argue that psychopaths do not share with us certain obviously true moral beliefs about the strength of our reasons to help and not harm others. In the next section, I argue that the truth of the Empathic Foundations Hypothesis provides the best explanation of why they fail to share our beliefs about our reasons to help others (section 3.1) and our reasons to avoid harming others (section 3.2). Section 4 addresses an objection from the armchair.

If true, the master argument establishes only that empathic concern at least *sometimes* provides a basic source of evidence regarding the strength of our reasons to help and not harm others. This conclusion leaves wide open the question of just *when* empathic concern provides this evidence. In section 5, I provide a partial account of the conditions under which empathic concern provides this evidence.

**Section 1: The Master Argument**

Most of us believe that we have strong reasons to help other people in need. Not just close friends and family, but strangers as well. We believe our reasons to help friends and family might be stronger than our reasons to help strangers, but we believe we have strong reasons to

help strangers all the same. We believe we have strong reasons to relieve them of suffering and deprivation, and to prevent their premature deaths.

We also believe we have strong reasons not to cause suffering and deprivation. We believe we have strong reasons not to harm innocent people. Some take our reasons not to harm to be stronger than our reasons to help, but nearly all of us take ourselves to have strong reasons of both kinds all the same. The beliefs that we have these strong reasons to help and not harm are among our most deeply held normative beliefs. Call the beliefs that we have these strong reasons to help and not harm our *core moral beliefs*.

I will argue that it would not be rational for us to retain our core moral beliefs if empathic concern were not a basic source of evidence regarding our pro tanto reasons to help and not harm. If it is rational for us to retain our core beliefs, it must be because there are at least some conditions under which experiences of empathic concern constitute a basic source of evidence regarding the strength of our pro tanto reasons to help and not harm.

I argue for this claim by reductio. The argument begins by granting for the sake of argument that:

(1) Empathic concern never constitutes a basic source of evidence for our core moral beliefs.

The next premise in the argument is an eminently plausible epistemic principle:

(2) If empathic concern never constitutes a basic source of evidence for our core moral beliefs, and we have no other evidence sufficient to justify our core moral beliefs, then we are rationally required to relinquish our core moral beliefs.

And now for a crucial premise:

(3) Excluding empathic concern, we have no other evidence sufficient to justify our core moral beliefs.

From these three premises we derive:

(4) We are rationally required to relinquish our core moral beliefs.

But (4) is clearly false:

(5) We are not rationally required to relinquish our core moral beliefs.

So we should reject our original assumption that empathic concern never constitutes a basic source of evidence for our core moral beliefs.

Four initial notes on the argument, all of which concern premise (1). First, by "a basic source of evidence" I mean a source of evidence that is not evidence in virtue of our having other kinds of evidence. For example, a basic source of evidence is not like a gas gauge, which can

provide evidence about the amount of gas in the tank only because we have independent evidence that the position of the gas gauge tracks the amount of gas in the tank.

Second, when I say that empathic constitutes a basic source of evidence for our core *moral beliefs*, I mean that it is a basic source of evidence regarding the strength of the *pro tanto* reasons we have to help and not harm, the reasons whose strengths give rise to the all-things-considered obligations that are the objects of our core moral beliefs.

Third, when I say that the evidence in question is evidence of the strength of *pro tanto* reasons to help and not harm, I mean that it provides evidence of how the strength of these reasons compares to the strength of *other* conflicting moral reasons, as well as how the strength of these reasons compares to the strength of our reasons to avoid personal sacrifice. These two different ways of assessing the strength of a reason correspond to Parfit's distinction between two different ways of assessing the strength of a duty. According to Parfit, duties can be stronger in the *conflict-of-duty* sense or duties can be stronger in the *cost-requiring* sense, where a duty is stronger than another in the conflict-of-duty sense just in case we are required to fulfill it rather than the other when we cannot fulfill both, and where a duty is stronger in the *cost-requiring sense* just in case we are required to undergo a greater personal sacrifice to fulfill it than we are required to undergo to fulfill the other (2017, p. 369) Although a reason could in principle be stronger in one sense without being stronger in the other, I will assume in what follows that they typically travel together and that empathic concern gives us evidence regarding the strength of our reasons in *both* senses.

Fourth, here and throughout the rest of the paper "evidence" refers to *non-testimonial* evidence. We may have plenty of testimonial evidence that our core moral beliefs are true, but the argument only considers scenarios where we lack such evidence.

With these clarifications out of the way, we can now ask: should we accept this argument? I take all but the third premise of this argument to be uncontroversial. The second premise is just an instance of the general epistemic principle that we're rationally required to relinquish our beliefs for which we lack sufficient evidence. And the assumption that (4) is false is obviously true; it appears to most of us that we are *not* rationally required to relinquish our core moral beliefs, and we have no reason to doubt that things are as they appear. But the third premise flies in the face of the now-dominant rationalist moral epistemologies, according to which our moral beliefs are justified because they *seem* true, or because they are *self-evident*. Moreover, this premise implies that someone who had never felt empathic concern in their life could know all of the morally relevant non-moral facts about torturing an infant for fun and yet still lack evidence sufficient to justify belief that it is wrong to torture an infant for fun. This seems quite wrong.

Nevertheless, I will argue that premise (3) is true. Again, I will argue by reductio. Suppose for the sake of argument that (3) is false and:

(6) Our non-empathic evidence is sufficient to justify our core moral beliefs.

Now consider an eminently plausible prediction:

(7) If our non-empathic evidence is sufficient to justify our core moral beliefs and there are people who (i) don't treat empathic concern as evidence, (ii) share our non-empathic evidence, and (iii) have no impairment in their ability to reason

from this non-empathic evidence, then these people should be just as likely as us to have our core moral beliefs.

From these two premises we derive:

(8) People who (i) don't treat empathic concern as evidence, (ii) share our non-empathic evidence, and (iii) have no impairment in their ability to reason from this non-empathic evidence should be just as likely as us to have our core moral beliefs.

But now for a crucial empirical premise:

(9) People who (i) don't treat empathic concern as evidence, (ii) share our non-empathic evidence, and (iii) have no impairment in their ability to reason from this non-empathic evidence are much less likely than us to have our core moral beliefs.

So we should reject our original assumption that we can lack empathic concern and yet have evidence sufficient to justify our core moral beliefs. We should accept (3).

The crucial premise in this second argument is premise (9). Why should we endorse premise (9)? Because we have sufficient reason to accept what I call *the Empathic Foundations Hypothesis* (EFH), and the Empathic Foundations Hypothesis implies that (9) is true. According to the EFH, we have our core moral beliefs only because we, or our testimonial informants, have

regularly felt empathic concern and treated it as evidence regarding the strength of our (pro tanto) reasons to help and not harm others. It follows from the EFH that, unless they rely on moral testimony, people who have not regularly felt empathic concern and treated it as evidence regarding the strength of their reasons to help and not harm others should not share our core moral beliefs—even if they have the same non-empathic evidence and have no impairment in their ability to reason from this non-empathic evidence. (For views that also assign empathic concern a central role in moral development, see Nichols 2004 and Blair 2017.)

Our rejection of premise (1) depends on our defense of premise (3). And our defense of premise (3) depends on our defense of premise (9), which in turn depends on a defense of the EFH. So the master argument that empathic concern constitutes evidence bearing on the strength of our reasons to help and not harm comes down to how well we can defend the EFH. In what follows, I provide an extended argument for the EFH. I argue that the EFH provides the best explanation of why we have our core moral beliefs but psychopaths don't, and that this is a good explanation.

In the next section, I argue that psychopaths do not share our core beliefs about our reasons to help and not harm. At the beginning of section 3, I articulate several predictions generated by the EFH. In section 3.1, I draw on these predictions to argue that the EFH provides the best explanation of why psychopaths don't share our core beliefs about helping: they feel less empathic concern than others and fail to take what little empathic concern they do feel as evidence regarding the strength of their reasons to help others. In section 3.2, I draw on these predictions to argue that the EFH provides the best explanation for why psychopaths don't share our core beliefs about harming: they feel less empathic concern than others and fail to take what

little empathic concern they do feel as evidence regarding the strength of their reasons to not harm others.

**Section 2: Psychopaths Do Not Share Our Core Beliefs**

We have a great deal of evidence that psychopaths do not share our core beliefs. One strong piece of evidence is that psychopaths do not act in the ways we would expect people to if they shared our core beliefs. First, they do not help others in need. Recent studies have found that psychopaths and those higher in psychopathic traits help those in need at a cost to themselves significantly less than neurotypical individuals (Sakai et al. 2012; 2016; 2017; Lockwood et al. 2017), especially when no one else is there to witness whether they help (White 2014). Second, they harm people with little hesitation whenever they believe that harming them will get them what they want. The latter tendency explains why psychopaths make up 15 to 25% of the North American male prison population despite making up only 1% or so of the North American male general population (Kiehl & Hoffman 2011). On the assumption that people typically act as they believe they have most reason to, this would provide strong evidence that psychopaths do not share our core beliefs, even if they claimed to.

But often they don't even claim to. Psychopaths explicitly deny that how an act impacts people's various rights and interests has an important role to play in determining whether it is right or wrong (Aharoni 2011; Blair 1995). The more psychopathic traits people have, the more likely they are to endorse (1) rational egoism (glossed as the view that "An action isn't rational if it doesn't aim to promote one's own self-interest"), (2) ethical egoism (glossed as the view that "An action isn't morally right if it doesn't aim to promote one's own self interest"), and (3) self-

serving, uncontroversially morally wrong business practices that harm others (Kahane et al. 2015; see also Decety & Yoder 2015). Highly psychopathic individuals also show far weaker endorsement of claims like "Compassion for those who are suffering is the most crucial virtue," "One of the worst things a person could do is hurt a defenseless animal," and "It can never be right to kill a human being." (Glenn 2009a) While these last three claims are not themselves among our core beliefs, they are closely related to our core beliefs, such that a decreased endorsement of these claims suggests a decreased endorsement of our core beliefs.

In some studies, psychopaths answer questionnaires in ways that make it seem like they share our core moral beliefs (Cima et al. 2010; Aharoni 2012; Aharoni et al. 2014; Marshall et al. 2017; Marshall et al. 2016). But it's likely that psychopaths in these studies are merely paying "lip-service" to the norms they hear people endorsing in their community. Indeed, when psychopaths are asked about less common kinds of moral questions—the kinds of questions they have not received explicit instructions about—their reactions deviate much further from the reactions of others. For example, numerous studies now suggest that those higher in psychopathic traits tend to think it is acceptable to kill one innocent person to save five innocent people (e.g., by pushing them in front of an oncoming train) (Bartels and Pizarro 2011; Koenigs et al. 2011; Seara-Cardoso 2013; Gao and Tang 2013; Djeriouat and Trémolière 2014; Miller et al. 2014; Wiech et al. 2013; see also Ritchie & Forth 2016), and those that fail to find such a correlation nevertheless find a strong correlation between psychopathic traits and individuals' *willingness* to kill one to save five themselves (Vyas et al. 2017; Tassy et al. 2013; Pletti et al. 2016; Cima et al. 2010; Seara-Cardoso 2012; Glenn et al. 2009a). Psychopaths also fail to condemn acts that cause fear (Cardinale and Marsh 2015; Marsh and Cardinale 2012; 2014) and accidental harm as strongly as non-psychopaths (Young et al. 2012). The fact that psychopaths

consistently answer these *unusual* moral questions differently than neurotypical individuals strongly suggest that they do not share our core moral beliefs. On the occasions where they seem to answer more familiar moral questions in the same way as others, this is because they know what other people want them to believe, not that they genuinely share our core moral beliefs.

One could object that in fact psychopaths do take our reasons to avoid harm seriously, and that whatever explains why they endorse the 'utilitarian' response to sacrificial dilemmas is whatever explains why many philosophers do. It could be objected that they do not condemn sacrificing the one to save the five in sacrificial dilemmas because they exercise their rational capacities to see that our reasons to avoid harm by omission are just as strong as our reasons to avoid harm by commission. But is this plausible?

It's true that psychopaths seem to assign equal weight to our reasons to avoid harming by commission and our reasons to avoid harming by omission. This could be because they believe that both types of reasons are very weighty—which would be compatible with their taking our reasons to avoid harm seriously—or it could be because they believe both types of reasons do not carry much weight at all. If it were because they believe both types of reasons carry little weight, this would not be consistent with the claim that they take our reasons to avoid harm seriously. And in fact, there is overwhelming evidence that this second explanation is the correct one. As already discussed above, psychopaths do not take our reasons to help seriously, and reasons to help are just reasons to avoid harm by omission. Moreover, as also mentioned above, Kahane finds that endorsing utilitarian judgments in the sacrificial dilemmas is strongly correlated with endorsement of ethical and especially rational egoism, as well as increased acceptance of self-serving business practices that harm others. This clearly reflects a failure to take our reasons to avoid harming others by omission seriously. So it seems that psychopaths' utilitarian judgments

about sacrificial dilemmas reflect a failure to take our reasons to avoid harm by commission seriously enough, rather than a tendency to take our reasons to avoid harm by omission especially seriously.

This all constitutes substantial evidence that psychopaths do not share our core beliefs. But the Empathic Foundations Hypothesis holds that they lack our core beliefs *because* they lack empathic concern or else fail to treat it as evidence bearing on the strength of their reasons to help and not harm others. Why should we endorse this stronger claim? It's uncontroversial that psychopaths lack empathic concern. But there are many psychological differences between psychopaths and neurotypical individuals. Psychopaths exhibit glibness, superficial charm, a grandiose sense of self-worth, and lack of remorse, among other things (Frick 1998; Hare 1999; Blair 2007). Why not think that one of these other differences could explain why psychopaths fail to share the core beliefs of neurotypical individuals?

An initial answer comes from the fact that although Glenn (2009a) found a negative correlation between how much participants exhibited psychopathic traits and how strongly they endorsed certain moral judgments about helping and harming, this correlation vanished once they controlled for differences in trait empathic concern. (*Trait empathic concern* refers to a person's tendency to feel empathic concern in response to certain types of stimuli.) In other words, after accounting for people's tendency to feel empathic concern, how much they exhibited the symptoms of psychopathy did not explain any differences in their moral judgments. A plausible explanation for this finding is that psychopaths do not share our core moral beliefs *because* they feel less empathic concern.

But the deeper answer is that we have good reason to endorse the Empathic Foundations Hypothesis, and the Empathic Foundations Hypothesis provides the best explanation of why neurotypical individuals endorse—and why psychopaths fail to endorse—our core moral beliefs.

## Section 3: Psychopaths Do Not Share Our Core Beliefs Because the Empathic Foundations Hypothesis is True

The Empathic Foundations Hypothesis provides an elegant story about how neurotypical individuals arrive at our core moral beliefs. Over the course of our moral development, we encounter various types of individuals enduring various types of misfortunes. When we encounter these individuals, we often feel empathic concern for them. We treat our empathic concern as providing us with evidence that we have reason of a certain strength to help or not harm them. As a result, we form the belief that we have reason of a certain strength to help or not harm others in certain situations. Over time, we generalize from these particular cases and come to form our core moral beliefs: that we generally have strong reason to help and not harm others in the ways described earlier.

The Empathic Foundations Hypothesis also provides an elegant explanation of why psychopaths fail to share our core moral beliefs. They fail to share our core moral beliefs because they lack the empathic concern that serves as the initial input to the process that results in our core moral beliefs. They lack evidence that they have reason to help and not harm others in particular cases. And to the extent they have any empathic concern in such cases, they fail to treat it as evidence regarding the strength of their reasons to help and not harm the individuals that triggered it. So whenever they encounter someone who has suffered a misfortune, they fail

to believe they have reason to help and not harm that person. Since they never see that they have strong reason to help and not harm others in any particular case, they fail to generalize to form our core moral beliefs. They fail to believe that we generally have strong reasons to help and not harm others.

This story makes a number of testable predictions. First, the EFH makes predictions about how levels of empathic concern should influence moral judgments about particular cases. We would expect for people to believe they have stronger reason to help and not harm someone in a particular case when they feel more empathic concern for them.

Second, not only should we find correlations between empathic concern and what people believe about their reasons, but we should also be able to increase how much reason people think they have to help and not harm an individual by causing them to feel greater empathic concern for that individual.

Third, we would expect that causing people to feel empathic concern would lead them to think they have stronger reason to help or not harm someone only if they had the capacity to interpret it as evidence and exercised that capacity. If people do not have the capacity to interpret their empathic concern as evidence or else do not exercise that capacity (e.g., because they have reason to doubt their empathic concern provides evidence in a particular case or they are under cognitive load), then they should be less likely to believe they have any reason to help.

Fourth, if the EFH is correct, we would expect people's trait empathic concern (i.e., how strongly disposed a person is to feel empathic concern) to correlate with how much they endorse our core moral beliefs. This is because such people will generally feel lower amounts of empathic concern during their moral development, and so they will tend to think their reasons to help and not harm in particular cases are weaker than those who feel more empathic concern.

This will in turn lead them to think their reasons to help and not harm are generally weaker. We have already seen that people with drastically reduced levels of empathic concern (i.e., psychopaths) do not share our core moral beliefs. But if the EFH is correct, endorsement of our core moral beliefs should also correlate with levels of trait empathic concern amongst those who are not psychopaths.

Fifth, the EFH makes a crucial developmental prediction: those who lose the ability to feel empathic concern following a normal course of moral development should share our core moral beliefs, but those who lose the ability to feel empathic concern early in their moral development should not. Those who lose the ability to feel empathic concern late in their moral development should share our core moral beliefs because they have accumulated a large body of evidence for those beliefs over the course of their moral development, whereas those who lose the ability to feel empathic concern early in their moral development have been deprived of this evidence.

I submit that if all five of these predictions are borne out, then the truth of the EFH provides an excellent explanation of why we have our core moral beliefs and psychopaths do not. Moreover, I submit that the EFH provides the *best* explanation of this fact. So we should endorse the EFH. In what follows, I'll provide evidence that all five of these predictions are borne out both in the case of our core moral beliefs about helping and our core moral beliefs about not harming. Since the empirical literatures on helping and harming are largely distinct, I will discuss evidence about helping and evidence about harming separately.

**Section 3.1: The Empathic Foundations Hypothesis Explains Why We Form Our Core Moral Beliefs About Helping and Psychopaths Don't**

The first prediction made by the EFH is that the amount of empathic concern people feel for a particular individual should correlate with how strong they take their reason to help that individual to be. This prediction receives strong support from numerous studies showing that how much people report *feeling* empathic concern in response to learning about someone else's misfortune strongly predicts how much they will help that person at some cost to themselves (Ashar et al 2017; Tusche et al. 2016; Hein 2010; Light et al. 2015; Barraza & Zak 2009; Barraza et al. 2015; Bornemann et al. 2016; Batson et al. 1981; Toi & Batson 1982; Fultz et al. 1986; Krebs 1975; see also FeldmanHall 2015; Morelli et al. 2014; Patil et al. 2017; Chopik et al. 2017; Davis 1994). If we can read off people's normative beliefs from their helping behavior, these data show that the amount of empathic concern felt towards an individual is correlated with how strong people take their reasons to help that individual to be.

While these data are impressive, they are merely correlational. If we want to be convinced that differences in empathic concern are not merely correlated with but are *causing* differences in judgment, then we need to turn from correlational data to experimental data. Turning to experimental data allows us to assess the second prediction of the EFH: that intervening to increase the amount of empathic concern people feel for an individual will increase how strong they take their reason to help that individual to be.

Decades of experimental interventions meant to increase how much empathic concern people feel for others—such as asking people to think about the thoughts and feelings of someone in need, having people look at the photograph of someone in need, etc.—have proven

to be effective ways of increasing people's willingness to help these individuals at some cost to themselves. Moreover, the effects of these interventions on helping are mediated by differences in how much empathic concern people report feeling for these individuals. (See Batson 2011 for a review.) Again, if we can read people's beliefs about the strength of their reasons from their behavior, this provides compelling evidence that empathic concern has a causal role to play in helping us see the strength of our reasons to help others.

Although this experimental data is compelling, it remains a live possibility that we have the causal arrow backwards. Instead of changing moral judgment by increasing empathic concern, it may be that these interventions first change moral judgment, which then changes empathic concern. If so, then these studies lend no support to the EFH. To rule this out, we need to find manipulations whose effects on empathic concern could not plausibly be preceded by differences in moral judgment.

Unfortunately, it is hard to be completely confident that any way of altering empathic concern itself does not first alter moral judgment. But there is another way to test whether empathic concern has a causal role to play in helping us see the strength of our reasons to help others. If the EFH is true and empathic concern typically changes moral judgment because people take their feelings of empathic concern for someone as *good evidence* that they have reason to help that person, then we can test whether empathic concern has a causal role to play in moral judgment by giving people reason to doubt their empathic concern constitutes evidence of this sort. If increasing empathic concern normally increases helping because people interpret it as good evidence that they have reason to help, then we should see that increasing empathic concern in such cases does not increase how much they help them. And if we see that increasing empathic concern in such cases does not increase helping, then this would provide compelling

evidence for both the second and third prediction of the EFH: that increasing empathic concern normally plays a causal role in helping us see the strength of our reasons to help others, and whether it plays this causal role depends on whether people have and exercise their capacity to interpret it as evidence.

So what we need is some way to convince people that the typical evidential relationship empathic concern bears to reasons to help has been severed. One way to do this is to convince people of some proposition whose truth would explain why they would feel empathic concern *even if* the proposition empathic concern was evidence for (i.e., that they had strong reason to help the empathic target) were false. This method consists in providing people with *undercutting defeaters* for any evidence that might be provided by empathic concern (Pollock 1986, p. 39).

While no studies relating empathic concern directly to moral judgment have been conducted using this methodology, several studies relating empathic concern to costly helping behavior have been. For example, Coke et al. (1978) found that when participants who were induced to feel empathic concern were convinced that their feelings were the product of a pill, empathic concern no longer increased the amount of help they provided to the person they felt empathic concern for. They helped no more than if they had never felt empathic concern for that person at all. The best explanation of why is that being given evidence that their feelings were the product of a pill convinced them that their empathic concern was not tracking the strength of their reasons to help. They believed they would have felt that empathic concern even if they had no reason to help at all. So the fact that they feel empathic concern provides no good evidence that they have reason to help. Using similar manipulations, Batson et al. (1981) and Stocks (2001) also found that giving participants convincing evidence that the source of their empathic

concern was something other than the misfortune of the target individual reduced levels of helping back to levels that are seen when no empathic concern is felt at all.

The studies just cited influenced whether people took their empathic concern as evidence by providing people with undercutting defeaters. But we can also influence whether people take their empathic concern as evidence by interfering with their ability to exercise their capacity to interpret empathic concern as evidence. For example, we can *distract* them. Hiraoka and Nomura (2016) found that putting people under cognitive load (i.e., having them memorize random strings of letters) decreased the correlation between empathic concern and intentions to help crying infants. That is, even if people feel just as much empathic concern when they are under cognitive load as when they are not, they are less likely to want to help. Hiraoka and Nomura (2017) found a similar effect looking at trait empathic concern rather than state empathic concern. A natural explanation of these findings is that exercising our capacity to interpret empathic concern as evidence takes a good deal of cognitive resources, and if our cognitive resources are occupied with something else—for example, counting—we cannot put them to use interpreting new evidence.

These findings lend strong support to predictions #2 and #3 of the EFH. They lend support to the view that empathic concern causes people to think they have reason to help the person who triggered their empathic concern, and that this is true *because* people treat their empathic concern as evidence that they have reason to help this person. But to vindicate the EFH, it isn't enough to show that people interpret particular episodes of empathic concern for particular individuals as evidence regarding the strength of their reasons to help those individuals. We must show that these particular episodes of empathic concern influence people's tendency to endorse our core moral beliefs about the strength of our reasons to help others

*generally*. If these particular episodes of empathic concern influence people's tendency to endorse our core moral beliefs, then we would expect two predictions to be borne out. These are the fourth and fifth predictions mentioned above.

According to prediction #4, trait empathic concern should correlate with endorsement of our core beliefs about helping in the general population. And it does. As already mentioned, while Glenn (2009a) found a negative correlation between how much participants in the general population exhibited psychopathic traits and how strongly they endorsed certain moral judgments about helping and not harming, this correlation vanished once they controlled for differences in trait empathic concern. In other words, after accounting for people's tendency to feel empathic concern, variation in psychopathic traits in the general population did not influence endorsement of core moral beliefs regarding helping. Trait empathic concern does all of the work explaining variability in abstract endorsement of core moral beliefs about helping. Another study has also found in the general population a strong relationship between empathic concern and endorsement of our core beliefs about the importance of helping those in need (Graham et al. 2011).

According to prediction #5, whether people endorse our core moral beliefs about helping should be affected by *when* they lose the ability to feel empathic concern. People who have lesions to their ventromedial prefrontal cortex (vmPFC) lack empathic concern (Damasio et al. 1990; Anderson et al. 1999; Barrash et al. 2000; Anderson et al. 2006; Eslinger et al. 2004). If prediction #5 is true, whether they endorse our core moral beliefs about helping should depend on when they incur their lesion, and hence when they lose their capacity for empathic concern. And it does. Patients who suffered damage to the vmPFC *late* in life share our core moral beliefs about our reasons to help others (Andersen et al. 1999; Saver & Damasio 1991), but patients with

early onset vmPFC damage do *not* share our core moral beliefs about the strength of our reasons to help others (Anderson et al. 1999). When given Kohlberg's famous moral dilemma of Heinz, in which Heinz has to choose between stealing a drug for his dying wife or letting his wife die, patients with late-onset vmPFC damage recognize the various opposing reasons at play—the reasons to avoid stealing and to help one's wife. But patients with early onset vmPFC damage reason in an egocentric manner, considering only how stealing or not stealing will effect the agent (i.e., by leading to punishment by authorities for stealing or by upset family members for not stealing). This reveals that they do not share our core beliefs about the strength of our reasons to help. While our core beliefs do not include beliefs about how Heinz should act, all things considered, our core beliefs do include beliefs about the nature and strength of Heinz's reasons to help his wife. We believe that Heinz has a strong reason to help his wife *for her own sake*, and not just because it will help Heinz avoid punishment. While late-onset vmPFC patients are sensitive to these kinds of reasons to help, these kinds of reasons are not even on the radar for early onset vmPFC patients.

These findings vindicate one half of the Empathic Foundations Hypothesis: we form our core moral beliefs about *helping* only because we feel empathic concern and treat it as a basic source of evidence regarding the strength of our reasons to help. Psychopaths don't form their core beliefs about helping because they don't feel as much empathic concern (Frick 1998; Hare 1999; Blair 2007), and we'll see in the next section that they have trouble treating what little empathic concern they do feel as evidence regarding the strength of their reasons.

**Section 3.2: The Empathic Foundations Hypothesis Explains Why We Form Our Core Moral Beliefs About Harming and Psychopaths Don't**

Recall that the first prediction of the EFH is that how much empathic concern someone feels for someone else predicts how strong they take their reasons to avoid harming that person to be. On the plausible assumption that people's behavior typically reflects their beliefs, we can assess our first prediction by seeing whether we find a strong correlation between levels of empathic concern and willingness to engage in unjustified aggression. Are people who feel less empathic concern for an individual more willing to harm them?

We can first investigate this question by looking at trait empathic concern. Do people who are disposed to feel empathic concern (i.e., have greater trait empathic concern) take their reasons to avoid harm to be stronger? Alas, much of the evidence on whether trait empathic concern has a strong relationship to unjustified aggression is mixed (see Eisenberg et al. 2010; Lovett and Sheffield 2007; Zych et al. 2017; van Langen et al. 2014; Jolliffe and Farrington 2004). Some studies suggest it does (Miller and Eisenberg, 1988), while others suggest it doesn't (Vachon et al. 2014). But most of the studies influencing the relationship between empathic concern and unjustified aggression rely on self-report, which is notoriously susceptible to bias, especially when certain answers are more socially desirable than others (Robinson and Rogers 2015). Fortunately, a recent study measured trait empathic concern in a way that gets around such bias, by asking about empathic responses to the suffering of particular individuals and then relating that to their general tendency to engage in aggression. As the Empathic Foundations Hypothesis would predict, it was found that participants who engaged in more aggression in general felt less intense empathic responses to these individuals (Winter et al. 2017). Another

recent study found that the only psychopathic traits that predicted how much aggression they engaged in was Callous Affect, a measure closely related to trait empathic concern (Lishner et al. 2015).

These studies all use trait empathic concern as a proxy for how much state empathic concern people feel for their potential victims. While this is a reasonable proxy, better yet would be to have laboratory studies in which state empathic concern is more directly measured, and in which people are given the opportunity to harm others. Unfortunately, studies of this sort are hard to come by. As it turns out, Institutional Review Boards are reluctant to approve experimental studies in which participants are given the opportunity to harm others. That being said, a recent study showing that how badly participants feel in response to shocking someone in exchange for money significantly predicts how likely they are to shock them again in the future (FeldmanHall et al. 2015). So we do have some recent evidence that a certain kind of empathic concern felt for an individual predicts willingness to harm that individual.

All of the studies discussed so far investigated people's willingness to harm by investigating their actual tendency to engage in harmful behavior. Another way to investigate willingness to harm people is by *asking* people whether harming others is morally permissible, or whether they would be willing to harm others under certain circumstances. Fortunately, there are now numerous studies asking people about their attitudes towards actively sacrificing one person's life (e.g., by throwing them in front of a train) to save five people's lives. If Prediction #1 were true, then we would expect those lower in empathic concern to be more approving of and more willing to sacrifice the one to save five in such cases. And many of the studies cited in section 2 on the relationship between psychopathic traits and moral judgment showed just this.

Even so, this evidence is merely correlational, so it only supports prediction #1. Do we have any data showing that experimentally increasing empathic concern reduces harmful behavior? If so, this would vindicate prediction #2. Again, IRBs do not often approve studies in which people are given (or believe they are given) the opportunity to harm others, so such studies are hard to come by. But there is one infamous exception to this rule: Stanley Milgram's experiments on obedience, in which participants were instructed by more or less authoritative experimenters to administer shocks to a confederate (the "learner") in the course of what was ostensibly a study on learning. Milgram did not measure empathic concern himself, but he did manipulate several factors known to influence whether people feel empathic concern for a victim: namely, what kinds of evidence they are given of the victim's suffering. As a recent meta-analysis shows, there was a strong relationship between the kinds of evidence participants were given and whether participants would administer these shocks (Haslam 2014). Participants were less willing to engage in this unjustified aggression when the victim was closer to them, less willing still when they heard the victim scream, and least willing when they had to physically put the learner's hand on a shock plate. One powerful explanation for why we see these effects on unjustified aggression is that each type of evidence, respectively, caused participants to feel more empathic concern, and this increased empathic concern prevented them from engaging in unjustified aggression. If this is right, then it supports prediction #2.

Although IRBs do not like to allow people to harm or to believe they are harming others, they are more accepting of asking people whether it's moral to harm others, or whether they would be willing to harm others under certain circumstances. And we have already seen that there are many IRB-approved studies asking people about their attitudes toward sacrificial dilemmas. Prediction #2 implies that if we could increase psychopaths' empathic concern for the

one in sacrificial dilemmas, this would decrease their willingness to sacrifice the one themselves as well as increase their condemnation of those who are willing to sacrifice the one. Existing evidence suggests this is the case. We see that in situations where people likely feel equally high empathic concern for everyone concerned—when everyone whose lives are at stake are friends or family members—they express significantly greater willingness to sacrifice the one to save the five (Kurzban et al. 2014). Likewise, in a virtual reality environment in which people can see and hear the five victims beckon for help (and are thus more likely to feel empathic concern for them), they are more likely to sacrifice the one (Patil et al. 2013; Francis 2016). Lastly, boosting serotonin with the use of citalopram (a selective serotonin reuptake inhibitor) caused people to more strongly condemn sacrificing one person, and this effect was higher amongst those who are high in trait empathic concern, suggesting that boosting serotonin changed moral judgment by boosting empathic concern for the one (Crockett et al. 2010).

These findings so far suggest that increasing someone's empathic concern for a victim increases the strength of the reason people take themselves to have to avoid harming that victim. But the EFH implies that we see this causal relationship specifically because people are interpreting their empathic concern to be *evidence* regarding the strength of their reasons to avoid harming the victim. To support this further claim, we need evidence for prediction #3: that empathic concern only increases how much reason people take themselves to avoid harming others when they have the capacity to interpret it as evidence and exercise that capacity. Unlike in the case of helping, no studies have been done in which people are convinced the empathic concern they feel is the product of a pill, or in which they are put under cognitive load. So we cannot find support for prediction #3 by interfering with people's ability to exercise their

capacity to interpret empathic concern as evidence. But we can find support for prediction #3 by looking at individual variability in people's ability to interpret empathic concern as evidence.

One way we can do this is by looking at those with alexithymia—a condition characterized by an inability to identify, describe, and rely on their feelings in reasoning—and those with psychopathy (Brewer et al. 2016). And as it turns out, those who are higher in alexithymia are more likely to endorse sacrificing the one to save the five in sacrificial dilemmas (Patil et al. 2016; Patil & Silani 2014b). They are also less likely to condemn both accidental (Patil & Silani 2014a) and intentional harms (Brewer et al. 2015). Since alexithymia has more to do with people's compromised ability to interpret their emotions rather than a compromised ability to feel emotions, these effects seem to be driven not by a failure on the part of alexithymic individuals to feel empathic concern for the person who has been harmed, but by a failure to interpret the empathic concern they do feel as evidence regarding the strength of the reason not to harm them.

Another way we can look for support for prediction #3 is by looking again at psychopaths. We now have excellent evidence that psychopaths not only feel less empathic concern in response to other people's distress (Aniskiewicz 1979; House & Milligan 1976; Blair 1999; Blair 1997; Seara-Cardoso et al. 2012; 2013 2016; Lockwood et al. 2013; Lee and Gibbons 2017; Lishner et al. 2015; Oliver et al. 2016; Caes et al. 2012; Pfabigan 2015; Decety et al. 2015; but see Lishner et al. 2012; Domes et al. 2014), but they also have trouble learning from the empathic concern they do manage to experience. Indeed, psychopaths *generally* have trouble using affect to learn from their mistakes, even when those mistakes harm only themselves (White et al. 2013; Finger et al. 2011; Fisher & Blair 1998; Obrien & Frick 1996; Blair et al. 20001; Budhani et al. 2006; Budahni & Blair 2005; De Brito et al. 2013; Fairchild et al. 2009; Newman

& Kossom 1986). This general problem with various types of reinforcement learning extends to the interpersonal domain; psychopaths fail to learn to avoid harmful actions not only because they feel less empathic concern, but because they fail to use what little empathic concern they do feel as input to reinforcement learning. (See Blair 2013 and Blair 2017 for reviews of the relevant literature.) This failure to use empathic concern as input to reinforcement learning just is a failure to treat the empathic concern they feel for a victim as evidence regarding the strength of the reason to have avoided harming the victim.

Let us move on to our final two predictions. Prediction #4 implies not only that those who feel minimal empathic concern (i.e., psychopaths) would fail to share our core moral beliefs about harming. It implies that trait empathic concern should correlate with endorsement of these beliefs even in the general population. Does it? Yes. As mentioned twice before, Glenn (2009a) found a negative correlation between how much participants in the general population exhibited psychopathic traits and how strongly they endorsed certain moral judgments about helping and harming, but this correlation vanished once they controlled for differences in trait empathic concern. And another study has also found in the general population a strong relationship between empathic concern and endorsement of our core beliefs about the importance of not harming others (Graham et al. 2011).

Finally, let us consider our fifth prediction. Prediction #5 implies that timing of empathic impairment should influence whether people accept our core beliefs about the strength of our reasons to avoid harm. And it does. People who have lesions to their ventromedial prefrontal cortex (vmPFC) lack empathic concern (Damasio et al. 1990; Anderson et al. 1999; Barrash et al. 2000; Anderson et al. 2006; Eslinger et al. 2004). But whether they endorse our core moral beliefs depends on when they incur their lesion, and hence when they lose their capacity for

empathic concern. Patients who suffered damage to the vmPFC *late* in life share our core moral beliefs about our reasons to avoid harming others (Taber-Thomas et al. 2014; Saver & Damasio 1991). Patients with early onset vmPFC damage do *not* share these core moral beliefs (Taber-Thomas et al. 2014; Anderson et al. 1999).

This concludes the argument for the second half of the Empathic Foundations Hypothesis. People endorse our core moral beliefs about not harming others only because they have felt empathic concern for those who have been harmed and treated that empathic concern as evidence that they have strong reason not to harm them.

***

Let's take stock. I have argued for the truth of the EFH. If the EFH is true, then premise (9) of our master argument is true: people who (i) don't treat empathic concern as evidence, (ii) share our non-empathic evidence, and (iii) have no impairment in their ability to reason from this non-empathic evidence are much less likely than us to have our core moral beliefs. Together with premises (7) and (8), this implies that premise (6) is false. But if premise (6) is false, then premise (3) is true: our non-empathic evidence is *not* sufficient to justify our core moral beliefs. Combining this insight with premise (1), the assumption we made for reductio—that empathic concern never constitutes a basic source of evidence regarding the strength of our reasons to help and not harm others—and premise (2)—the uncontroversial idea that we are rationally required to relinquish the belief that P when we lack evidence sufficient to justify believing P—we can derive the conclusion that we are rationally required to relinquish our core moral beliefs. But (5) is clearly true: we are *not* rationally required to relinquish our core moral beliefs. So we should

reject premise (1), the assumption we made for reductio: it's *not* the case that empathic concern never constitutes a basic source of evidence regarding the strength of our reasons to help and not harm others.

**Section 4: An Objection from the Armchair**

I have argued that it is rational for us to retain our core moral beliefs only if empathic concern at least *sometimes* provides us with evidence regarding the strength of our reasons to help and not harm. If it *never* provided us with evidence regarding the strength of our reasons to help and not harm, then we would be rationally required to relinquish our core moral beliefs. Since it is clearly irrational for us to reject our core beliefs, empathic concern must provide us with evidence regarding our reasons to help and not harm under at least some conditions.

At this point, the empathy skeptic might object that the story I have told does not accurately describe why *they*, the skeptics, share our core moral beliefs. They may hold that they share our core moral beliefs because the Light of Reason has revealed these beliefs to be true in a way that does not implicate empathic concern. They may hold that rational reflection alone provides them with evidence sufficient to justify forming their core beliefs. In other words, even if premise (3) is true of most people, it is not true of them. If that's right, then the rationality of retaining their core beliefs puts no pressure on them to accept that empathic concern *ever* provides good evidence regarding the strength of our reasons to help and not harm. They can rationally endorse their core beliefs while rationally denying that empathic concern ever provides good evidence regarding the strength of their reasons to help and not harm.

The skeptic's story enjoys some degree of plausibility. After all, if I asked you right now why it was rational for you hold your core moral beliefs, what would you say? You wouldn't say anything about feeling empathic concern for anyone. You would probably say that it just *seems* to you that these beliefs are true.

Nevertheless, we have excellent reason to believe they only have this seeming because they have felt empathic concern and treated it as evidence in the past. This may not be obvious. Indeed, I have just spent thousands of words arguing for this non-obvious conclusion. But we should endorse this non-obvious conclusion. And if this non-obvious conclusion is true, it's not rational for them to believe that their *seeming* provides evidence sufficient to justify their core moral beliefs while at the same rejecting the view that empathic concern provides evidence regarding the strength of their reasons to help and not harm. Doing so would be like trusting their hunch that Mexico will pay for the wall even after they recognize that they only have this hunch because they previously treated what was not an authoritative news source—Donald Trump's Twitter feed—as though it were. It would be to disavow the influence of some evidence while at the same time covertly relying on it. It would be irrational.

To resist this line of argument, the empathy skeptic would have to insist that their core beliefs really do not depend in any way on having ever treated empathic concern as evidence in the past, or at least that they can *now* rationally derive their core beliefs from an alternative source of evidence: their unadulterated capacity for rational reflection. Is this at all plausible?

If we could come to grasp the truth of our core beliefs via rational reflection, then we would expect that anyone who freely exercised this unimpaired capacity would also grasp the truth of these beliefs if they considered them, regardless of whether they could experience empathic concern. In support of this view, the empathy skeptic might offer the vmPFC patients

who lack empathic concern but who still share our core moral beliefs (Taber-Thomas et al. 2014; Anderson et al. 1999). But recall that it is only patients who suffered damage to the vmPFC *late* in life that share our core moral beliefs. Patients with early onset vmPFC damage do *not* share our core moral beliefs. If rational reflection were sufficient to enable people to see the truth of our core beliefs, then we would expect *all* vmPFC patients to endorse our core moral beliefs, absent some other epistemic disadvantage they might have. But we do not see this.

To resist this argument, the empathy skeptic would have to argue that when we knock out vmPFC, we also knock out whatever rational capacities allow us to see the truth of our core beliefs without the aid of empathic concern. But this is special pleading. While the vmPFC is involved in many tasks, it does not seem to be involved in the relevant sort of rational reflection. When people engage in cold rational reflection about moral (Greene et al. 2001; Glenn et al. 2009b; Kuehne et al. 2015) or non-moral content (Goel & Dolan 2003; Curtis & D'Esposito 2003; Barbey et al. 2013; Barbey et al. 2014), it is the dorsolateral prefrontal cortex (dlPFC)—not the vmPFC—that is engaged. And the dlPFC is intact and functioning normally, both in vmPFC patients and psychopaths. So we should reject the empathy skeptic's claim that their endorsement of our core moral beliefs does not depend on empathic concern. Those who have never experienced empathic concern but otherwise seem perfectly capable of rational reflection do not share our core beliefs, even when they explicitly consider them.

At this stage, the empathy skeptic might concede that empathic concern plays an important role in helping us form our core moral beliefs but deny that it plays the role of evidence. But if it didn't play the role of evidence, why would it play such an important role in our moral development? I will consider what I take to be the only two plausible proposals.

One proposal is that empathic concern is required for us to possess moral concepts. If this were true, then the problem with psychopaths is not that they lack evidence for our core moral beliefs, but they are not even talking about the same thing as us in the first place! This is not plausible. Despite the common suggestion that psychopaths lack moral concepts (Kennett & Fine 2008; Prinz 2007), there is little reason to believe this. (See also Kumar 2016, who argues that psychopaths have an incomplete grasp of moral concepts without lacking moral concepts altogether.)

Why might someone think that psychopaths lack moral concepts? First, one might think that in order to be competent with moral concepts, one must have at least some true moral beliefs about uncontroversial matters—they must share our core moral beliefs. It might be held that someone who did not share our core moral beliefs (e.g., a psychopath) could not possibly be talking about the same thing as us when they use the terms "morally right" and "morally wrong" (Prinz 2007, p. 43; Kennett & Fine 2008, p. 176). This assumption is highly controversial. But suppose it were right. Does this provide an adequate explanation of why we need empathic concern to arrive at our core moral beliefs that does not rely on the assumption that empathic concern plays the role of evidence in that process? No. The evidence reviewed above strongly suggests that the role empathic concern plays in producing our core moral beliefs is that of evidence. The proposal on offer adds that a failure to share our core moral beliefs prohibits someone from possession with moral concepts. But this does not compete with the explanation that empathic concern plays the role of evidence in generating our core moral beliefs. It depends on it.

Alternatively, one might hold that psychopaths must lack moral concepts because they perform poorly on Turiel's moral/conventional task (Prinz 2007, pp. 43-44; Kennett & Fine

2008, p. 174), in which participants are asked whether a host of prototypical moral violations (e.g., bullying) and conventional violations (e.g., wearing mismatched socks) are wrong, serious, authority independent (i.e., would be wrong even if it was not condemned by any authority), and general in scope (e.g., would be wrong in different times and places). Most people judge prototypical moral violations to be wrong, serious, authority-independent, and general in scope, but they do not judge prototypical conventional violations to have these features. (See Turiel 1983 for Nucci 2001 for reviews of the relevant literature.) Psychopaths, on the other hand, treat both moral and conventional violations similarly. Adult psychopaths seem to treat conventional violations the way that non-psychopaths treat moral violations (Blair 1995; Blair et al. 1995), and young psychopaths seem to treat moral violations the way that non-psychopaths treat conventional violations (Blair 1997). Neither pattern of judgment, however, suggests that psychopaths lack moral concepts. It merely shows they have different patterns of judgment, and, as argued in the previous paragraph, this does not plausibly show that psychopaths lack moral concepts.

It might be held instead that these results show that psychopaths do not possess concepts because they do not possess certain concepts one must possess in order to possess moral concepts: the concepts of seriousness, authority-independence, and generality. But this is far from obvious. If psychopaths did not understand these concepts, we would expect their ratings of seriousness, authority-independence, and generality to be more or less random guesses. But this is not what we see. There are systematic patterns in how both adult and youth psychopaths rate the seriousness, authority-independence, and generality of norm violations. This is not what we would expect if they did not possess these concepts.

Another way to argue that psychopaths lack moral concepts is to rely on the fact that they are not competent with moral concepts because they are not motivated to comply with the moral judgments they seem to make (Kennett & Fine 2008, pp. 176-177; also see Kumar 2016 for discussion). This argument relies on the assumption that the moral norms they report endorsing but violating are moral judgments they in fact have. I have already given substantial reasons to doubt that this is so. But even if we assume that psychopaths fail to be motivated by whatever moral judgments they do in fact have, this only delivers the result that they lack competence with moral concepts if we assume a strong and controversial form of judgment internalism, the view that making an occurrent moral judgment is necessarily accompanied by some motivation to comply with that moral judgment. That being said, even if we do assume that strong and controversial form of judgment internalism, it's far from obvious why being unmotivated to comply with their moral judgments should influence what moral judgments they make. And what we need here is an explanation of why psychopaths seem to make different moral judgments than the rest of us, not an explanation of why their (apparently moral) judgments are not *moral* after all. So even if it's true that psychopaths lack moral concepts, this cannot explain what needs to be explained: why they do not endorse the same rules as us, whether their endorsement counts as *moral* or not.

In any case, and more importantly, the main problem with any view that holds empathic concern is required only for us to gain moral concepts is that we have already seen compelling evidence that empathic concern plays the role of evidence in moral judgment and moral learning. And the view the empathic concern only contributes to the possession of moral concepts cannot explain these findings. The proposal that empathic concern merely plays a role in allowing us to possess moral concepts cannot explain why experiencing varying degrees of empathic concern in

particular cases influences moral judgment in the way it seems to. It is not as if people who feel more empathic concern in a particular case take themselves to have stronger reasons to help because they temporarily have a firmer possession of moral concepts. It is far more plausible to hold that they take themselves to have stronger reasons to help because the greater empathic concern provides them with evidence that they have stronger reasons to help.

A second proposal for why empathic concern might play an important role in moral development without providing evidence is that it is required to get people to pay enough *attention* to moral matters for them to learn about them. On this view, empathic concern simply functions as an alarm bell, getting people to pay attention to the non-empathic evidence that is sufficient to justify their core moral beliefs. While I do not deny that empathic concern can play this function, taking it to play *only* this function likewise leaves unexplained a number of findings already discussed (e.g., why amounts of empathic concern track the strength of the reasons people take themselves to have, why believing the empathic concern to be caused by a pill prevents it from having its normal influence on moral judgment). Moreover, if empathic concern were simply directing our attention, then we would expect its influence on moral judgment to be more switch-like: once we have enough empathic concern to direct our attention and get us engage, further empathic concern should not influence moral judgment. Surely attention of various forms (e.g., visual attention) comes in degrees, but it seems that the kind of attention in question (i.e., noticing that someone has suffered a misfortune) does not come in degrees. This is because the kinds of non-empathic evidence people think empathic concern directs our attention to (e.g., that someone has suffered a misfortune) seems to be the kind of thing we either notice or we don't. If that's right, then if empathic concern were only playing the role of directing attention, we should see a threshold of empathic concern above which we get

moral engagement and below which we do not. But again, the influence that empathic concern has on moral judgment is far more graded.

I have considered two alternative explanations of why empathic concern plays a crucial role in moral development. On the first view, empathic concern is solely required for the possession of moral concepts, without which there could be no moral development. On the second view, empathic concern is solely required to direct attention to the non-empathic evidence on which normal moral development depends. I have found both of these explanations wanting. Empathy skeptics should therefore conclude that they share our core moral beliefs only because they have, over the course of their moral development, treated their empathic concern as evidence for the truth of these beliefs just like everyone else who is not a psychopath. They should conclude further that rational reflection does not give them non-empathic evidence sufficient to justify their core beliefs. So if empathy skeptics don't want to commit themselves to relinquishing their core beliefs, even they must accept, by their own lights, that empathic concern at least *sometimes* provides good evidence regarding the strength of their reasons to help and not harm.

But suppose I am wrong, and rational reflection *does* provide us with non-empathic evidence sufficient to justify our core moral beliefs. If this were true, then empathy skeptics could deny that empathic concern ever provides good evidence for our core moral beliefs without committing themselves to thinking that they are rationally required to relinquish their core moral beliefs. Nevertheless, they would remain committed to something almost as bad. That is, they would be committed to denying that most people are doxastically justified in holding their core moral beliefs. They would be committed to this because the studies cited in support of the EFH strongly suggest that most people form their core moral beliefs only because they treat their

empathic concern as evidence in support of those beliefs. And since one cannot be doxastically justified in believing P if one believes P only because one treats E as evidence for P when E is not in fact evidence for P, most people would not be doxastically justified in holding their core moral beliefs if empathic concern were not evidence for those beliefs. But most people clearly are doxastically justified in holding their core moral beliefs. It appears clear to most of us that we are so justified before we engage in any rational reflection, and we have no reason to doubt that things are as they appear. It follows from this that empathic concern must at least sometimes be evidence of the sort I say it is.

Moreover, if rational reflection does provide an alternative route to our core moral beliefs, this does not give us any reason to doubt that empathic concern provides evidence of the sort I say it does. This is because empathic concern *also* leads us to form our core moral beliefs. By the lights of rational reflection, empathic concern seems not to be leading us astray. Indeed, if we think that rational reflection gets us to the same beliefs that relying on empathic concern as evidence does, this seems to provide an independent check on the reliability of empathic concern. So we should embrace the view the empathic concern at least sometimes provides evidence of the sort I say it does.

**Section 5: *When* Does Empathic Concern Constitute Good Evidence?**

There are many factors that influence how much empathic concern we feel for others. Under certain conditions, we might feel a great deal of empathic concern for someone. Under other conditions, we might feel very little empathic concern for that very same person. If feeling more empathic concern for someone provides us with evidence that our reasons to help that

person are stronger, empathic concern seems to be telling us different things in different conditions. Sometimes this is how things should be, since our reasons to help the person might be much stronger under some conditions than others. For example, it makes sense that we feel more empathic concern for someone when they have suffered a grave misfortune, such as losing a loved one, than when they have suffered a minor misfortune, such as stubbing their toe. But other times, it seems empathic concern tells us different things in different conditions even though we have good reason to think the strength of our reasons has not changed across these conditions. For example, we feel much more concern for people when we see their face and hear their voice. But merely seeing someone's face and hearing their voice does not change our reasons to help someone. Nor does it inform us of any reasons to help them that we might have been neglecting. The strength of our reason to help someone does not turn on the shape of their face or the pitch of their voice. So empathic concern seems to tell us different things in different circumstances, even though the facts it purports to be telling us about have not changed.

It seems empathic concern is sensitive to changes that it shouldn't be. Skeptics like Paul Bloom, Jesse Prinz, and Peter Singer take this as decisive reason to reject the evidential value of empathic concern wholesale (Bloom 2016; Prinz 2011; Singer 1974). But following them would be hasty. It would be like rejecting the evidential value of visual experience based on the fact that we sometimes dream. And in any case, we have seen that we cannot completely reject the evidential value of empathic concern on pain of having to relinquish our core beliefs. So the way to respond to the fact the empathic concern is sensitive to morally irrelevant factors is not to dismiss the testimony of empathic concern altogether. It is to figure out when we should trust it and when we shouldn't.

In setting up the problem above, I suggested that empathic concern provides conflicting testimony in different conditions. But this isn't quite right. Even if empathic concern tells us different things in different conditions, it does not follow that it is telling us incompatible things. Consider that when we take our glasses off, we might see a blurry figure. This will give us some evidence regarding the shape of the object in front of us. When we put on our glasses, we get different evidence regarding the shape of the object in front of us. But our visual experiences in these cases are not delivering conflicting testimony. Our visual experiences give us evidence in one case for a more determinate version of what they give us evidence for in the other. We see the same phenomenon in the case of empathic concern. This is because empathic concern provides us with evidence regarding the *minimal* strength of our reasons to help and not harm. So though we might feel more empathic concern for someone when we see their face as compared to when we don't, it does not follow that empathic concern is providing us with conflicting testimony in these two cases. The small amount of empathic concern provides us with evidence that we have at least a weak reason to help—leaving open the possibility that our reason to help might be even stronger—while the moderate amount of empathic concern provides us with evidence that we have at least a moderately weighty reason to help. Having at least a moderately weighty reason to help is a more determinate version of having at least a weak reason to help. Evidence for the one does not conflict with evidence for the other.

So empathic concern might constitute good evidence across a variety of conditions even when the morally relevant facts have not changed across conditions and even when we feel different levels of empathic concern across those conditions. Nevertheless, it provides us with evidence for more determinate propositions in some conditions (i.e., that the minimal strength of our reasons to help or not harm is higher), and having evidence for such determinate propositions

can make a significant difference to what it's rational for us to believe about how we should act. So we need to ask whether we should trust empathic concern under the conditions that increase how much of it we feel, and thereby increase the determinacy of the proposition (i.e., the strength of the reason) it seems to provide evidence for.

I have already mentioned two of the seemingly irrelevant conditions that increase how much empathic concern we feel for an empathic target: seeing the target's face and hearing their voice. There are others: (accurately or inaccurately) perceiving the target's face to be cute (Batson et al. 2005; Lishner et al. 2008; Zickfeld et al. 2018), perceiving the target's face (Västjfäll et al. 2014; Kogut and Ritov 2005a), accurately perceiving the target's voice to be expressing sadness, fear, or pain (Phillips et. al 1998; Lishner et al. 2008; Lang et al. 2011; Quas et al. 2017), thinking vividly about the thoughts and feelings of the target (Batson 2011), (truly or falsely) believing the target to be close by, and having interacted with the target in the past in certain positive ways (Hein et al. 2016; Hein et al. 2010; Johnston and Glasford 2017; Batson and Ahmad 2009; Pettigrew and Tropp 2008). The conditions in question all typically increase empathic concern when they obtain. But whether these conditions obtain does not bear on the strength of our reasons to help the target. Nor must these conditions give us access to any morally relevant non-moral considerations that do bear on the strength of our reasons. (In some cases they may, but they increase empathic concern even in cases where they do not.)

Should we be willing to rely on empathic concern under these conditions, just as we should be willing to rely on our visual experiences when we open our eyes and turn on the lights? Or should we be unwilling to rely on empathic concern under these conditions, just as we should be unwilling to rely on our visual experiences when we close our eyes and take hallucinogenic drugs?

Each of these conditions has an independent influence on how much empathic concern we feel, so we need to ask for each condition whether it undermines the credibility of empathic concern. But I will consider five of the conditions—seeing the target's face, seeing fear or sadness or pain in their face, hearing their voice, hearing fear or sadness or pain in their voice, and thinking vividly about their feelings—as a package, since they typically travel together. I will then consider perceived cuteness, perceived proximity, and past interaction.

Should we be willing to trust empathic concern when we see the target's face, see fear sadness or pain in their face, hear their voice, hear fear sadness or pain in their voice, and think vividly about their feelings? Yes. We should think that when these conditions obtain, it's more like we're opening our eyes and turning on the lights than closing our eyes and downing some hallucinogenic drugs. Why must we think this? Because if empathic concern were untrustworthy under these conditions, then we would have to think that, all else being equal, we would have been better moral judges if we had never found ourselves in these conditions. But, as I will shortly argue, if we had never found ourselves in these conditions, all else being equal, we would have felt more or less no empathic concern at all. And we have already seen that those who felt more or less no empathic concern at all would not have come to share our core moral beliefs. Since someone who fails to share our core moral beliefs cannot be a better moral judge than we are, we should reject the assumption that we should not trust the deliverances of empathic concern under the conditions in question.

The contentious premise in this argument is that if we had never found ourselves in these conditions, we would have felt more or less no empathic concern at all. Why think this? First, it's a plain fact that we don't often feel empathic concern for strangers when we're not in these conditions. This is why it isn't enough for fundraising charities to merely *tell* people that there

are some needy individuals who could use their help. They have to show them slow-motion videos and tell them vivid stories describing the plight of these individuals. Second, there is now excellent evidence that psychopaths' failure to feel empathic concern is in large part explained by an inability to spontaneously put themselves in some of these favorable conditions. In particular, it is explained by an inability to spontaneously perceive fear or sadness in other people's faces, voices, or postures (see Blair 2017 for a review), as well as a tendency to not think vividly about other people's feelings (Meffert et al. 2013; Lui et al. 2016; Beussink et al. 2017). This deficit seems to be responsible for psychopaths' reduced empathic concern and subsequent antisocial behavior. In striking support of this claim, a recent study found that improving the ability of young criminal offenders (and likely psychopaths) to recognize fear and sadness in other people's faces reduces recidivism (Hubble et al. 2015; see Dadds et al. 2012 for a similar study). Other studies find that if psychopaths and those who are high in psychopathic traits deliberately think vividly about other people's feelings (e.g., by imaginatively putting themselves in the other person's shoes), they can experience increased empathic concern (Lui et al. 2016; Meffert et al. 2013; Arbuckle & Shane 2017; Beussink et al. 2017). These findings suggest that psychopaths are not constitutionally incapable of feeling empathic concern. They just have trouble getting themselves into some of the circumstances that trigger empathic concern.

So if we want to know how much empathic concern we would feel if we never found ourselves in these five conditions, we need only look once again at psychopaths. When we do, we see we would feel very little empathic concern. And we've already seen that those who feel little empathic concern do not share our core moral beliefs. If empathic concern felt under these five conditions was in fact untrustworthy, then we should think (all else being equal) that those

who did not feel empathic concern under these conditions would be better moral judges. But they clearly aren't better moral judges. They're psychopaths. So we should reject the assumption that we can't trust empathic concern under these conditions.

The argument thus far establishes only that empathic concern is more trustworthy when the five conditions listed above are met than when none of them are met. But this may be true even if some of the conditions reduce the credibility of empathic concern, so long as the credibility deficit is made up for the by the fact that other conditions increase its credibility even more. But in order for this to be the case, some of these conditions must be increasing empathic concern while others decrease it. And we know that each of these conditions increases empathic concern. So if each increases empathic concern individually, and the increased empathic concern they produce collectively is trustworthy, the increased empathic concern they individually contribute must be trustworthy as well.

It remains to be shown that three additional conditions—perceived cuteness, perceived proximity, and previous interaction—do not undermine the credibility of empathic concern. To establish that none of these conditions undermine the credibility of empathic concern, we would have to show that people do not make worse moral judgments when each condition obtains compared to when each condition does not obtain, holding fixed the morally relevant non-moral facts. I will show this by arguing that people tend to make *better* moral judgments when each condition obtains.

Why think we're better judges of our reasons to help or not harm someone when we perceive their face to be cute, when we believe them to be close by, and when we've interacted with them in the past? Because we do not care too much about cute individuals; we care too little about individuals who are not cute. Parents tend to be less motivated to care for and be

affectionate towards their infants when their infants are less cute (Kringelbach et al. 2016; Langlois et al. 1995; Glocker et al. 2009). It does not seem that those who care for and are affectionate towards their infant in the normal ways are overestimating the strength of their reasons to help their children in these ways. Likewise, we do not care too much about those who are close by or who we have interacted with in the past; we care too little about individuals that are far away and with whom we have had no positive interactions. Most of us tend to care very little about those who are far away or with whom we have had no positive interactions. It is implausible to think that our obligations to those who are nearby and with whom we have had positive interactions are similarly weak (Singer 1972; Cf. Kamm 2000).

Empathic concern under conditions of cuteness, proximity, and previous interaction constitutes better evidence than empathic concern in outside of these conditions, just as the visual experiences we have when wearing corrective lenses constitute better evidence than the visual experiences we have when we're not bespectacled. Saying this is not to endorse preferential treatment for cute individuals, closeby individuals, or those we have interacted with in the past. It is to say that we would have the highest quality evidence about the strength of our reasons to help a non-cute, non-proximal individual with whom we have never interacted if we imagined and felt empathic concern for a cute, proximal version of that individual with whom we have previously interacted. If we cannot do this, then we should take the empathic concern we feel for a cute, nearby individual with whom we have interacted (e.g., a dog) not only as evidence regarding the strength of our reasons to help that individual, but as evidence regarding the strength of our reasons to help non-cute, non-proximal individuals with whom we have never interacted when those individuals are the same in all morally relevant respects (e.g., a pig).

**Section 6: Conclusion**

In this chapter, I argued that empathic concern at least sometimes provides a basic source of evidence regarding the strength of our reasons to help and not harm others. I argued that if this were false, we would be rationally required to relinquish our core moral beliefs—or at least that most of us would not be doxastically justified in holding our core moral beliefs. Since we are not rationally required to relinquish our core moral beliefs, and most of us are doxastically justified in holding our core moral beliefs, empathic concern must at least sometimes provide a basic source of evidence regarding the strength of our reasons to help and not harm others. I argued further that empathic concern continues to provide good evidence—and in fact provides higher quality evidence—under some surprising conditions: when we perceive the target's face to be cute, when we accurately perceive the target's face to be expressing sadness, fear, or pain, when we accurately perceive the target's voice to be expressing sadness, fear, or pain, when we think vividly about the thoughts and feelings of the target, when we believe the target to be close by, and when we have have previously interacted with the target in certain ways. In the next chapter of the dissertation, I provide an account of when we should expect the evidence provided by empathic concern to improve moral judgment, all things considered. In the final two chapters of the dissertation, I take lessons from chapter 1 and chapter 2 and use them to resolve debates regarding the strength of our reasons to prevent animal suffering and to give to charities fighting extreme poverty.

## Chapter 2

## How to Collect Empathic Evidence

In the previous chapter, I argued that empathic concern provides a basic source of evidence regarding the strength of our reasons to help and not harm others. I argued further that it provides high quality evidence under a variety of conditions, with it providing evidence for more determinate propositions under some conditions and less determinate propositions under others. But even if empathic concern does provide high quality evidence under all of these conditions, it does not immediately follow that obtaining this evidence always improves moral judgment. As opponents of empathic concern frequently point out, empathic concern has well known limits. We feel empathic concern for our near and dear to the exclusion of those perceived to be distant or strange. One individual in need drums up far more empathic concern than even two or three. Most people's reservoir of empathic concern dries up long before they can feel the appropriate amount of empathic concern for everyone who deserves it. And the amount of empathic concern we feel for an individual often "maxes out" even when our reasons to help and not harm that individual remain quite modest in strength. There's a case to be made that, in light of these limits, collecting the evidence provided by empathic concern will often hinder more than it will help us in figuring out how we should act. It will provide us with one epistemic advantage, but only at the cost of introducing even greater epistemic disadvantages.

In light of its limits, how should we go about collecting the evidence provided by empathic concern so that it improves moral judgment? What is the *proper* way to collect the evidence provided by empathic concern? Under what conditions would we expect collecting the evidence provided empathic concern to corrupt moral judgment more than it corrects it?

To arrive at an answer to this question, I will begin in the first half of the chapter by considering exactly how the limits of empathic concern can undermine its epistemic utility. Once we have such an explanation, we can then specify a set of conditions under which the facts that constitute the explanation do not obtain, and thus in which the limits of empathic concern do not undermine its epistemic utility.

Even if we know when the limits of empathic concern fail to undermine its epistemic utility—when the limits are irrelevant—we may not yet know how to acquaint ourselves with the evidence provided by empathic concern in a way that will improve moral judgment. In particular, it may not be that feeling empathic concern under just *any* conditions in which the limits are irrelevant will improve moral judgment. It may be there are other conditions—conditions unrelated to overcoming the previously identified limits of empathic concern—that must be met in order for empathic concern to improve moral judgment. It may be that it's not enough to feel empathic concern on one occasion, or whenever it's convenient. Indeed, I will argue in the second half of this chapter, I will offer several views on additional conditions that must be met in order to collect the evidence provided by empathic concern in such a way as to improve moral judgment, noting their relative advantages and disadvantages and the contexts in which abiding by some views is better than abiding by others.

**Section 1: Limit #1: The Closeness Bias**

Perhaps the most salient limit of empathic concern is that we tend to feel it in full force only for our nearest and dearest. For example, we feel empathic concern most for people who are close to us in time and in space, who are members of our ingroup, and to whom we have some

sort of attachment. Let us call this limit the *closeness* bias. Under what conditions does the closeness bias undermine the epistemic utility of empathic concern, and why?

The fact that empathic concern exhibits this kind of bias undermines the epistemic utility of feeling empathic concern only in cases where it (a) leads us to have less accurate beliefs about the absolute strength of our reasons to help or not harm someone, or (b) leads us to have less accurate beliefs about the *relative* strength of our reasons to help or not harm multiple individuals.

There are two ways we can have inaccurate beliefs about the absolute strengths of our reasons to help and not harm others. One way is to overestimate their strength. The other way is to underestimate the strength. So we must ask: are we overestimating the absolute strength of our reasons to help and not harm our nearest and dearest? Or are we underestimating the absolute strength of our reasons to help and not harm others? It's plausible to think that some of the factors that make people near and dear to us—e.g., the fact that we are friends or family—do increase the strength of our reasons to help (and perhaps not harm) them. So to the extent these factors increase the amount of empathic concern we feel for our near and dear, there is little reason to think they are giving us less accurate beliefs about the strength of the reasons we have to help and not harm them. But other factors mentioned in the previous chapter—mere proximity, etc.—do not increase the strength of our reasons. Nevertheless, I argued in that chapter that the presence of these additional factors does not undermine the quality of evidence provided by empathic concern. Indeed, it increases the quality evidence provided by empathic concern, as it allows empathic concern to provide us with credible evidence of a more determinate proposition than it would if these factors did not obtain. These considerations suggest that this initial limit of empathic concern—that we feel it most fully only for our nearest

and dearest—does not undermine its epistemic utility by leading us to overestimate the strength of our reasons to help our nearest and dearest.

If the closeness bias does not undermine the epistemic utility of empathic concern by leading us to overestimate the strength of our reasons to help our nearest and dearest, then we can say this much already: *when only the interests of our nearest and dearest are relevant to a moral question, the closeness bias does not undermine its epistemic utility.*

But might the closeness bias undermine the epistemic utility of empathic concern when the interests of those who are not our nearest and dearest are at stake? There are two ways this could happen. One is for empathic concern to lead us to *underestimate* the absolute strength of reasons to help and not harm them. The other is for empathic concern to lead us to overestimate the relative strength of our reasons to help our nearest and dearest compared to our reasons to help others.

I have already argued that empathic concern gives us evidence that our reasons to help and not harm others achieve *at least* a certain strength. If this is right, the diminished empathic concern we feel for those who are not our nearest and dearest is not directly misleading us into thinking our reasons to help them are especially weak. Just as being unable to see a distant star with our naked eye does not rebut any independent evidence we might have (e.g., through astrophysics) that there is such a star, feeling too little empathic concern for an individual does not rebut any independent evidence we might have (e.g., through rational reflection) that our reasons are even stronger than empathic concern suggests. If we had independent evidence that our reasons to help these people were exceptionally strong, we would not interpret the low levels of empathic concern we feel for them as evidence that our reasons to help them are weaker than we thought. So the closeness bias does not undermine the epistemic utility of empathic concern

by leading us to underestimate the absolute strength of our reasons to help and not harm those who are not our nearest and dearest.

It follows that the only way the closeness bias could undermine the epistemic utility of empathic concern is by leading us to have inaccurate beliefs about the *relative* strength of our reasons to help our nearest and dearest and those who are not our nearest and dearest. It's clear to see how it could do this. It could give us especially strong evidence that our reason to help our nearest and dearest are as strong as they in fact are. But it could also *fail* to provide us with strong evidence that our reasons to help others are as strong as they in fact are. As a result, allowing ourselves to feel empathic concern freely could lead us to think our reasons to help our nearest and dearest are much stronger than our reasons to help others, not because empathic concern gives us evidence that we have only weak reason to help those who are not near and dear, but because it *fails* to give us evidence that our reasons to help those who are not near and dear are as strong as they in fact are.

It is quite plausible that the closeness bias could undermine the epistemic utility of empathic concern for these reasons. So this should lead us to think that, compared to feeling no empathic concern at all for any of the individuals whose interests are stake in a particular moral question, feeling empathic concern in a way that is infected by the closeness bias may harm moral judgment more than it helps it. This is not to say that it will *always* harm moral judgment more than it helps it. If feeling no empathic concern left us without any evidence that we had reasons to help and not harm *anyone*, then that would be worse than a case in which it gave us evidence that we had stronger reasons to help and not harm our nearest and dearest compared to others. So what this suggests is that we should be on the look out for the closeness bias when collecting the evidence provided by empathic concern. *In cases where we the closeness bias is*

*not leading us to form inaccurate beliefs about the relative strength of our reasons to help and not harm close others compared to non-close others, the closeness bias does not undermine its epistemic utility*. We would expect such cases to obtain when we make those who are not near and dear to us "close"—by interacting with them in virtual reality, looking at their faces, etc.—in the same way that our near and dear to us are.

I should note before moving on that the closeness bias cannot only lead us to mistakenly think that our reasons to help our near and dear are much stronger than our reasons to help others, but it can lead us to mistakenly think that our reasons to help one individual in the *short-term* is much stronger than our reasons to help them in the *long-term*, because the stages of the person in the near future are much closer to us in the relevant respects than the stages of the person in the further future. We can draw the same conclusion as before: *in cases where the closeness bias is not leading us to form inaccurate beliefs about the relative strength of our reasons to help and not harm a person in the short-term compared to the long-term, the closeness bias does not undermine its epistemic utility*. We would expect such cases to obtain when we make the person's future interests salient and thus "close" in the way their short-term interests are.

**Section 2: Limit #2: Innumeracy**

Critics of empathic concern frequently point out that empathic concern is "innumerate"—we do not typically feel more empathic concern for a group of individuals in need than we would if we focused on just one of those individuals. On the plausible assumption that we have stronger reason to help multiple individuals in need than just one of those individuals, this provides an

explanation of how the innumeracy of empathic concern could undermine moral judgment: it could lead us to think that our reasons to help a group of individuals is no stronger than our reason to help just one individual. We can draw the following conclusion: *in cases where the innumeracy of empathic concern does not lead us to think that our reasons to help a group of individuals is no stronger than our reason to help just one individual, its innumeracy does not undermine its epistemic utility*.

One class of cases where we would expect innumeracy not to undermine the epistemic utility of empathic concern are cases in which the only interests at stake are the interests of the agent/moral inquirer and one other individual. For example, as I will discuss in the next chapter, one such case is when we are deciding whether we have stronger reason to give ourselves the pleasure we get from eating an animal product than we do to avoid bringing about the suffering that animals had to endure in order for us to get that pleasure. Likewise, as I will discuss in the final chapter, another such case is when we are deciding whether we have stronger reason to give ourselves the benefits of some luxury purchase or to come to the aid of someone who would have been benefit had we instead of made that purchase donated an equivalent amount to a highly effective charity.

As I will argue in chapters 3 and 4, we may also overcome innumeracy in cases in which the interests of many individuals are stake, if we can break up these multi-party cases into a series of dyadic cases of the sort just discussed. For example, suppose we are trying to decide whether we ought to donate $10,000 to help 100 individuals. If we attempt to feel empathic concern for all 100 individuals simultaneously, this will not give us access to the strength of our reasons to help these 100 individuals. This is because empathic concern is innumerate. But we can break up this question into a series of smaller questions. We can ask whether we ought to

give $100 to one individual. Then we can ask whether we ought to give another $100 to another individual. And so on and so forth. If we were to carry out this procedure and decide in each small-scale case that we ought to give the $100 to each individual in each case, then we can reasonably infer that we ought to give all $10,000 to help all 100 individuals all at once in the large-scale case. If each step in the process of giving our $10,000 away is required and doing as we're required in each case leads us to giving all of our $10,000 away, then it's reasonable to assume that we would be required to bring about the same outcome directly, by giving all $10,000 away at once. (I will discuss objections to this strategy in chapters 3 and 4, where I will deploy it to argue that we ought to abstain from animal products and give away substantial amounts of money to charity.)

Alternatively, such cases may obtain when we can feel empathic concern in a *serial* fashion for each of the individuals whose interests are stake, rather than attempting to direct empathic concern directly at a group of individuals whose interests are bundled together by the nature of the choice we face. For example, I can empathize with my friend, and while empathizing I can make a judgment (perhaps unconsciously) about how strong my reason to promote their pleasure is. Next, I can empathize with the animal, and while empathizing I can make a judgment (perhaps unconsciously) about how strong my reason to prevent that suffering is. On the assumption that these judgments about pro tanto reasons can outlive the empathic states that generated them, I can later draw on them when making an all-things-considered judgment about how I am morally obligated to spend my money. In such cases, empathic concern remains innumerate in a given case, but we can overcome its innumeracy. (Although I believe this is a way of overcoming innumeracy in principle, I will express doubts below about our ability to successfully carry out this procedure in practice.)

**Section 3: Limit #3: Burnout**

Critics of empathic concern are also quick to point out that people who regularly feel empathic concern suffer from burnout. The stock example of burnout victims are nurses. Nurses are regularly exposed to the kinds of conditions that lead them to feel empathic concern, and as a result they begin to avoid such situations, to feel less empathic concern in those situations, or to be less moved to help by the empathic concern they feel in those situations. Some of these worries about burnout are worries about the motivational power of empathic concern. But we want to know how burnout could undermine the epistemic utility of feeling empathic concern. How could it?

The epistemic worry about burnout is that there may be cases in which we reach burnout before we can feel empathic concern equally for everyone whose interests are at stake. As before, this will lead us to think that our reason to help some individuals are stronger than our reason to help other individuals. This is just another case in which we see the same problems as with the closeness bias and innumeracy. It does not introduce any new problems. The solution, then, is quite similar to our previous solutions: *limit our use of empathic concern to cases where we do not achieve burnout*. In cases where some of us reach burnout before others, this will give us reason to defer to the judgments of those who can avoid burnout.

Of course, this solution to the problem of burnout generates a second worry: that when we stop feeling empathic concern due to burnout, this is not because we can no longer exercise our capacity for empathic concern due to brute physiological limits, but that we continue to

exercise that capacity, and there happen to be no reasons for it to detect. I address this worry in chapters 3 and 4.

**Section 4: Limit #4: Ceiling Effects**

The amount of empathic concern we feel for a particular individual does not always covary with the size of their misfortune. In particular, it often seems to be insensitive to facts about the magnitude of the target's misfortune and the certainty that they have suffered or will suffer that misfortune. For example, we do not always feel more empathic concern for a loved one who is dying at the age of 10 than a loved one who is dying at the age of 40, even when dying at 10 is a greater tragedy than dying at 40. The amount of empathic concern we can feel has already hit ceiling when we move from considering death at 40 to death at 10. Likewise, we do not feel more empathic concern for a loved one who has a 99% chance of dying tonight than a loved one who has a 90% chance of dying tonight. The amount of empathic concern we feel will likely have already hit ceiling long before the chance of dying reaches 90%. Since the strength of our reason to help someone depends in part on the size of their misfortune and the likelihood that they will suffer it, empathic concern should be sensitive to these relevant influences. But it is not always sensitive to it. I will offer three explanations for how the amount of empathic concern could be insensitive to the strength of our reasons but remain good evidence for the strength of those reasons.

**Section 4.1 Explanation #1**

One explanation for why empathic concern can be insensitive to the size and likelihood of someone's misfortune in these ways is that empathic concern has "limited bandwidth". Just as our visual experiences can only provide evidence about objects within our immediate vicinity, empathic concern can only provide evidence about reasons of a certain strength. It can only carry so much information about the strength of a reason. For reasons below a certain strength, the amount of empathic concern we feel will vary in accordance with the strength of that reason. But for reasons above that strength, the empathic concern we feel will be equally strong. So empathic concern cannot help us discover how these reasons differ in strength, and it may lead us to mistakenly think that our reasons to help or not harm are equally strong, even if they are not. In such cases, the insensitivity of empathic concern may rob it of its epistemic utility. It follows from these observations that *before relying on empathic concern, we must ensure we are not in a case where the reasons empathic concern is giving us access to exceed the limited bandwidth of empathic concern.*

**Section 4.2 Explanation #2**

Even so, the fact that empathic concern has limited bandwidth cannot explain all cases in which empathic concern is insensitive to the strength of our reasons, and so it will not always be enough for us to ensure we are not in a case where the limited bandwidth of empathic concern is irrelevant. For example, we might feel the same moderate level of empathic concern for a child who is being bullied on the playground and a child who, living in a war zone, cowers in fear at

night as bombs whizz overhead. Although our reasons to help both children are quite strong, it is clear that our reason to help prevent a child from being killed in war is stronger than our reason to prevent a child from being bullied. It might seem empathic concern could not be informing us of this difference in strengths, since we feel the same amount for both.

But this would only be so if equal amounts of empathic concern provided evidence for reasons of equal strengths across all contexts. And this may be false. Two phenomenally indistinguishable experiences of empathic concern may be capable of providing evidence for reasons of different strengths in different contexts. To see how this might work, consider a different case where two phenomenally indistinguishable pieces of evidence clearly provide evidence for different claims in different contexts. Suppose you are looking at two photographs taken from one end of a hallway, looking down the hallway. In the first photograph, a person appears to be standing at the far end of the hall. If you were to measure their height on the page, they would come out to be 2 cm tall. Now suppose the other photo is a photoshopped version of the first one where the 2 cm tall image of the person has been moved so it appears this person is now standing in the foreground, right in front of the camera. The parts of the photos that depict the person in both of these photographs are intrinsically phenomenally indistinguishable. (Or, if not intrinsically *phenomenally* indistinguishable, they at least impact our eyes in indistinguishable ways.) They take up the same amount of space on the page. More importantly, the light they reflect occupies just as much area on the retina, and they occupy just as much of an angle within one's visual field. So the intrinsic character of our experience of this person remains constant regardless of which photograph we are looking at. Nevertheless, our intrinsically identical experiences of the person provide very different evidence regarding the size of the person we are looking at. When the person appears in the background, our experience of the

person provides evidence that they are a normal height. But when the person appears in the foreground, our intrinsically identical experience of the person provides evidence that they are extremely small, a miniature version of a normally sized human. So we have a clear-cut case where two phenomenally indistinguishable pieces of evidence provide evidence for different claims in different contexts. I claim that something similar can happen in the case of empathic concern: feeling a certain level of empathic concern can provide evidence that we have a weak reason to help in one case, but feeling that same level of empathic concern can provide evidence that we have a strong reason to help in another case.

How could this be? How could feeling a certain level of empathic concern in one case provide evidence for a strong reason while that same level of empathic concern in another case provides evidence for a weak reason? To see how, first consider why it is that the 2 cm image of the person can provide evidence that they are tall in one context but provide evidence that they are very short in another. The reason is that these contexts differ in *depth cues*, the cues they provide about the distance between the person represented by the image and the observer. When the 2 cm image appears in the foreground of the hallway, we have evidence that the person represented by the image is close to us. When the 2 cm image appears far down the hallway, we have evidence that the person represented by the image is far away from us. Since we know that larger items appear smaller the farther away they are, this changes what conclusions we should draw about the person's height on the basis of our visual experiences of them. Phenomenally indistinguishable experiences of the person provide evidence that they have different heights in the two photos because depth cues differ in the two photos. The difference in depth cues allows us to make a correction for distance when consulting our evidence.

I have just argued that differences in depth cues allow two visual experiences that are intrinsically phenomenally indistinguishable to provide evidence for different claims in different contexts. If a given level of empathic concern can provide evidence of different reason strengths in different contexts, then there must be some difference between these contexts that can play the same role as depth cues. I submit that whenever equivalent levels of empathic concern provide evidence for reasons of different strengths in different contexts, these will be contexts in which we find differences in *psychological distance* between us and the targets of our empathic concern. A moderate amount of empathic concern can provide evidence for a strong reason to help when the target of our empathic concern is psychologically distant, but it can provide evidence for a smaller reason to help when the target of our empathic concern is psychologically near. (As I am thinking of it, how much psychological distance we have to a particular individual is matter of the extent to which the morally irrelevant relations identified in the previous chapter—the relations that increase empathic concern for an individual—obtain.)

Consider how differences in psychological distance explain the case of the bullied child and the child at war. Given how we represent the bullied child, they are likely to be represented as psychologically close. We imagine knowing the bullied child, we likely see their face and can imagine hearing their sobs. We imagine them on the playground on the school near where we live. If we have ever been bullied ourselves, we can vividly imagine their thoughts and feelings. But given our representation of the child at war, they are likely to be represented as psychologically distant. We will represent them as physically far away, so they will be spatially distant. They will likely be a member of some salient outgroup, speaking a different language, practicing a different religion, so they will be socially distant. And we might not be able to

vividly imagine their face or their sobs, or their thoughts and feelings. So our representation of them will be abstract.

Insofar as we recognize that the child at war is psychologically distant, we can take advantage of the evidence provided by empathic concern to improve moral judgment. Although our level of empathic concern is low compared to the strength of our reason to help, we can make a "correction" for psychological distance. Just as we do a correction for spatial distance when we interpret the image of the person at the end of the hallway as evidence that they are quite tall, we do a correction for psychological distance when we interpret our moderate empathic concern for the child at war as evidence that we have strong reason to help them. This is why a moderate level of empathic concern directed toward a child at war can provide evidence that we have extremely strong reason to help, whereas a moderate level of empathic concern directed toward a bullied child can provide evidence that we have quite strong but not as strong reason to help.

If this is all right, then a given level of empathic concern can provide evidence of reasons with different strengths in different contexts. This means that although our levels of empathic concern do not always change as the magnitude of the target's misfortune change, it does not follow that it is misleading us in either of these cases. So long as we can take into account differences in psychological distance, feeling empathic concern can aid moral judgment even when it seems to be insensitive to differences in the magnitude of the target's misfortune.

**Section 4.3 Explanation #3**

I have claimed that equal levels of empathic concern can provide evidence for reasons of different strengths only when psychological distance varies, and so the observed insensitivity of empathic concern to reasons of different strengths need not undermine its epistemic utility in such cases. But the explanation I have given of why the insensitivity of empathic concern does not undermine its epistemic utility depends on there being differences in psychological distance in the two episodes of empathic concern observed. Unfortunately, there seem to be cases where we feel equal levels of empathic concern under equal levels of psychological distance, and yet the misfortune that the empathic concern is a response to provides reasons of different strengths. For example, suppose we see a child who is currently suffering from the cold and we feel a moderate amount of empathic concern. Now suppose we see a child who is currently suffering from the same symptoms, but in fact these are symptoms of a much longer-lasting disease. We have stronger reason to help in the second case than we do in the first. But since we feel equal amounts of empathic concern in both cases, it seems that empathic concern could not be informing us of that fact. Feeling empathic concern in this case might lead us to mistakenly think that our reasons to help both children are equally strong.

I submit that in cases of this sort—where we are focused on the ongoing physical or emotional distress of the other person under conditions of minimal psychological distance—our empathic concern provides evidence regarding our reasons to prevent them from continuing to feel what they are currently feeling in the immediate future. Empathic concern in such cases is a response to a relatively short-lived misfortune—the person's imminent suffering—rather than the prolonged misfortune of which this imminent suffering is just a small part. It is fortunate that

empathic concern can narrow its focus in this way, because otherwise our empathic concern would constantly be "maxing out". If we were always focused on prolonged suffering, we would constantly be feeling the maximum amount of empathic concern, and thus everything would seem equally bad.

The situation we have here is much like the situation we have when photographing objects at close range. When objects are close to the camera, more of the light reflecting off the object makes its way through the camera's lens, and the photograph becomes overexposed. Everything ends up looking the same—maximally bright. One way to get around this is to reduce shutter speed—to allow the photograph to be based on a smaller duration of exposure to the object. In this way, we can reintroduce variation in the photographs of different objects.

Likewise, when targets of empathic concern are psychologically close to empathizers, more of the reasons to help the targets make their way to the empathizer, and the empathic concern the empathizer feels hits its maximum. Everything ends up looking the same to them—as if they have equally strong reason to help all targets. But just as we can reduce shutter speed to avoid overexposure, we can reduce the period of time our empathic concern is a reaction to. In this way, we can reintroduce variation in the strengths of the reasons empathic concern informs us of. Because even if the strength of the reason to prevent the total misfortune of each target surpasses the reason strength that empathic concern can inform us of, the reason to prevent their temporary suffering can vary depending on (e.g.) the intensity of their suffering.

I have considered three explanations of why empathic concern might be insensitive to the strength of the reasons we have to help and not harm others. First, the reasons in question might exceed the limited bandwidth of empathic concern. Second, there may be varying levels of distance between us and the individuals who elicited the empathic concern. Third, even when

there are equal levels of distance between the individuals in question, the empathic concern may only be a response to a portion of each individual's misfortune, portions that we have equally strong pro tanto reason to prevent, even if the total misfortune of one is worse than the total misfortune of the other. In light of these explanations, we can say what conditions must be met if the insensitivity of empathic concern is not to rob it of its epistemic utility. First, it must not be the case that more than one of the reasons at stake exceeds the limited bandwidth of empathic concern. Second, when the empathic concern we feel is felt under different conditions of psychological distance, we must retain and exercise the ability to "correct" for psychological distance. Third, when the empathic concern is directed toward a temporally limited misfortune, we must retain and exercise the ability to calculate how strong our reason to prevent the temporally extended misfortune is in light of the evidence empathic concern provides about the strength of our reason to prevent a temporally limited part of that misfortune.

It is not always easy to know when these conditions are met. But it is easy to know that these conditions are met in the cases I discuss in the rest of this dissertation. In the case of whether to consume animal products, the reasons at stake—i.e., the reason to give ourselves pleasure and to prevent animal suffering—do not both exceed the limited bandwidth of empathic concern, both individuals whose interests are at stake—i.e., ourselves and the animal—are equally psychologically distant (and so we need not worry about whether we are adequately exercising an ability to "correct" for psychological distance), and the interests in question—i.e., a moment of pleasure and a moment of suffering—are temporally limited (and so we need not worry about our ability to calculate the strength of our reason to prevent a prolonged misfortune on the basis of evidence regarding the strength of our reason to prevent a portion of that misfortune). In the case of forfeiting luxury goods to donate to charity,

the reasons at stake—i.e., to indulge in luxury goods or to prevent great suffering from extreme poverty—do not *both* exceed the limited bandwidth of empathic concern, and both individuals whose interests are at stake—i.e., ourselves and the person who we could help by donating to charity—are equally psychologically distant (and so we need not worry about properly exercising the ability to correct for psychological distance). And while the interests at stake—i.e., our *regular* indulgence in luxury goods throughout our lives, and the *prolonged* suffering or *significant* deprivation of those in extreme poverty—are not temporally limited, it is not plausible to think that the judgments we make on the basis of empathic concern in these cases reflects an inability to recognize the strength of the reasons we have to promote these prolonged interests, of which our empathic concern is a reaction to only a part.

**Section 5: Additional Conditions**

Suppose we are considering a moral question where we have good reason to think that none of the limits of empathic concern discussed above will rob empathic concern of its epistemic utility. In these cases, collecting the evidence provided by empathic concern will advantage us by giving us additional evidence, and it won't introduce any other epistemic disadvantage that outweighs the advantage provided by gathering this additional evidence. But are all ways of collecting this evidence equally advantageous? Are there certain conditions under which we can better *appreciate* the evidence provided by empathic concern, in which we respond to the evidence provided by empathic concern in a more rational manner?

There are several plausible views about when we should feel empathic concern in order to best incorporate it into our moral judgment. Suppose we want to know whether an action A is

permissible. On what we might call the *synchronic one-off* view, we should feel empathic concern for all of the individuals affected by A on one occasion, *while* we issue our moral judgment about A. On what we might call the *serial one-off* view, we should feel empathic concern for each of the individuals affected by A on one occasion, one at a time, immediately before we issue our moral judgment about A. Lastly, on what we might call the *habitual* view, we should feel empathic concern for individuals affected by *actions like A* every time we perform A-like actions, simulate performing A-like actions, or witness others performing A-like actions, at the time those actions occur. On this view, what is important is that we have felt empathic concern in the past for those affected by actions similar to the action A we are currently issuing a moral judgment about. We need to feel empathic concern for those who are harmed by the action, or those who could have been helped by the action. Like the serial one-off view, the habitual view holds that we need not feel empathic concern for those affected by A *while* we issue our moral judgment about A. Instead, it holds that we need to feel empathic concern for those affected by A-like actions whenever we perform, simulate, or observe others performing A-like actions.

Following the instructions of each of these views has certain advantages and disadvantages. In favor of the synchronic one-off view, it seems likely that we are best placed to respond rationally to our evidence while we possess it, and we only possess the evidence constituted by empathic concern while we are feeling empathic concern. Against the synchronic one-off view, we can note that the cases in which it will be useful are highly circumscribed, since we can only feel empathic concern for one individual at a time. Indeed, a large body of evidence reveals that people feel *less* empathic concern for multiple individuals in need than they do for a particular individual (Västfjäll et al. 2014; Kogut & Ritov 2005a; Kogut & Ritov 2005b; Jenni &

Loewenstein 1997). As a result, it seems the synchronic one-off view will only be of use in cases where the only interests at stake in whether an action is performed are the interests of the moral judge and the target of empathic concern.

Unlike the synchronic one-off view, the serial one-off view allows us to in principle use empathic concern to gather evidence about our reasons to perform actions that affect *multiple* individuals. Since the serial one-off view allows us to feel empathic concern for each of the individuals affected by the action in a serial fashion, that it means it allows for a delay between the episode of empathic concern we feel for an individual and the time at which we issue our moral judgment about the action that affects them. Against the serial one-off view, we might hold that we are not as well placed to exploit the evidence provided by empathic concern when we are not feeling it, just as (e.g.) we are not as well placed to evaluate the beauty of a painting when we are not looking at it, or to count the stars in the sky after we have closed our eyes. In order for the serial view to be useful, we need some reason to believe that we can exploit the evidence provided by empathic concern even when we no longer have it. The most natural way this could happen would be for us to use the empathic concern we feel for each individual *while we are feeling* it to arrive at a judgment about the strength of the reason we have to help or not harm that individual, and to then store that judgment for later use when issuing our all-things-considered moral judgment about A.

In order for the serial one-off view to be one we ought to follow, we need some reason to think we have the right cognitive architecture to carry out its instructions. In particular, we need the ability to create a representation of the outcomes of an action we are evaluating, to represent the value of those actions (or, alternatively, our reasons to bring about those actions), *and to update those values on the basis of empathic concern in a lasting way*. While it is well

established that we have these first two abilities—which go under the heading of *model-based decision-making* in recent computational neuroscience—it's an open question whether we have the last ability. While many of the studies discussed in chapter 1 provide strong evidence that we do use empathic concern to update our evaluative representations in particular decisions, it is unclear whether we can update these evaluative representations on the basis of a single episode of empathic concern. Some reason to think we do not is that even if people feel equal empathic concern for multiple individuals in a serial fashion, they tend not to treat their reasons to help each of those individuals equally. Instead, they tend to favor the final charitable target they consider (Huber et al. 2011). This suggests either that feeling empathic concern just once cannot change our mental models, or it changes them unevenly.

The habitual view gets around the worry that single episodes of empathic concern may not be powerful enough to update the representations we draw on in issuing moral judgments. The habitual view relies on our having the ability to update ongoing evaluative representations of certain act-types on the basis of the evidence provided by empathic concern and the ability to exploit these representations in online decisionmaking. These abilities correspond to what has been much discussed recently under the heading of *model-free* and *model-based* reinforcement learning and decision-making. For model-based reinforcement learning, we rely on an enduring representations of the possible outcomes of actions, probabilities attached to those outcomes, and values assigned to those outcomes. Model-free learning is so-called because it does not depend on any enduring representations of the outcomes of actions For model-free learning, we simply maintain cached values of certain action types, which we can update as the evidence provided by empathic concern comes streaming in. (See Crockett 2013 and Cushman 2013 for seminal papers introducing reinforcement learning to the topic of empirical moral psychology.)

The habitual view may be our best bet for incorporating all of the evidence provided by empathic concern. But the habitual view has certain disadvantages. Recall that the habitual view requires us to have repeatedly felt empathic concern for individuals affected by actions that are *similar to* the action whose moral status we are currently investigating. When following the habitual view engages *model-free* reinforcement learning, we arrive on the basis of this process at a *general* judgment about the moral status of A-like actions, and we then do a bit of deductive reasoning in particular cases to arrive at a judgment about the particular action A that we are presently evaluating. Importantly, whether collecting the evidence provided by empathic concern in the habitual way will aid us accurately evaluating the present action A depends in large part in how similar A is to the A-like actions in the past. If the A-like actions are sufficiently different than A, model-free learning may lead us to think that A has one moral status, when in fact the factors that made A-like actions in the past wrong do not obtain.

Consider one prominent example of this line of argument. Joshua Greene (2017) argues that we should not trust our aversion to pushing an individual off of a footbridge to stop a trolley from killing five innocent people, because *this* case of pushing an individual to their death is different from all of the other similar cases we have encountered in the past. In the past, instances of pushing individuals to their deaths almost always produced outcomes that were bad on the whole. They did not save people's lives. They only resulted in the death of one and the grief of others. Because the case of pushing the person off the footbridge is different from the cases our model-free learning judgment of the case is based on, Greene argues we should not trust this judgment. The factors in past cases that led us to believe such actions to be generally impermissible do not obtain in the present case, so we should not generalize from the past to the present, as we would be doing if we relied on model-free learning.

Regardless of whether Greene has provided the correct explanation of our aversion to pushing people off of footbridges, we can see how it would call into question our judgments about trolley cases if it were correct. This is all to simply point out that although the habitual view provides the best bet for incorporating the evidence provided by empathic concern, we must we wary of the possibility that following the habitual view may lead us to treat present cases like past cases, even when there are morally relevant differences between them.

The habitual view says that in order to exploit the evidence empathic concern provides for whether we should perform certain types of actions, we must feel empathic concern for individuals affected by those actions *during or shortly* after performing those actions. Why? Because the habitual view is helpful only because we have the capacity to use reinforcement learning. And in order for reinforcement learning to take empathic concern into account when generating evaluations of actions, the empathic concern has to be felt during or soon after the action is performed. Otherwise, these mechanisms will have trouble assigning the consequences that triggered the empathic concern to that action. It will be ambiguous which action the consequences are a product of, and thus it will be ambiguous what the evidence provided by empathic concern is evidence *for*. The way to minimize ambiguity is to ensure that the empathic concern is felt immediately after the action. So, I submit, we can better appreciate the evidence provided by empathic concern when it is felt immediately after the action whose normative status it provides evidence of. More concretely, to help children learn that it's wrong to hit others, you want to minimize the delay between their hitting others and their witnessing the consequences for the person they hit. Likewise for adults, if we want to learn whether treating another individual a certain way is wrong, we need to witness the consequences of that treatment

immediately and feel any empathic concern there is to feel. It is not enough for us to see the consequences on videotape later—although that would likely be better than nothing.

Moreover, it isn't enough for us to perform the action once, observe the consequences immediately, feel empathic concern, and then go on with our lives. The reinforcement learning processes that produce our judgments base their evaluations of different action types on the evidence received each time we perform actions of that type. So our judgments about whether certain types of acts are permissible are the products of the evidence we've gathered *across time*. This means that to figure out the deontic status of a certain type of act that we regularly perform, it isn't enough to collect the evidence provided by empathic concern on just one occasion. We have to collect it on as many occasions as possible—ideally, on every occasion where we perform that action or we witness someone else performing that action. If we ignore the evidence provided by empathic concern, we will be missing crucial evidence that we had strong reason to avoid acting as we did. Indeed, if the act brought us some kind of benefit, all we will have is evidence that we had reason to perform. So the body of evidence on which our judgments are based will be biased. And this biased body of evidence will lead us to falsely believe that actions of that sort are permissible.

What I just said may not apply in every case. In a case where we are performing a genuinely novel action—where we start our moral learning fresh, as it were—it may be enough to collect the evidence provided by empathic concern on that one occasion. This is because our judgment will not be unduly influenced by a biased set of evidence that we've consulted in the past. But most controversial moral questions are not questions about actions we have never performed before, or that bear no resemblance to actions we have performed before. One reason why they are controversial is that they are actions that many people perform and others believe to

be wrong. In these kinds of real-world cases, we *will* have performed many similar actions in the past. For example, in chapter 3, I will be discussing the action of purchasing and consuming animal products. Since most people have done this multiple times a day for most of their lives, it will not be enough for them to experience empathic concern for the animals harmed by their consumption just once. Having done this, the total set of evidence on which their judgment is based will remain severely biased. If they wanted to unbias their judgment, they will have to collect the evidence provided by empathic concern every time they perform the action in question.

These considerations also bear on when we should follow the instructions of one-off views. If we follow a one-off view and it leads us to revise our judgments regarding a particular action A, we should think that our new judgment is more likely to be correct than our previous judgment, as it is informed by the evidence provided by empathic concern. But if we follow a one-off view and we do not change our judgment, we should not rest easy. It may be that our judgment is the product of model-free reinforcement learning that is not itself properly informed by empathic concern. If we have reason to believe this, then we should defer to the judgments we have decisive reason to believe we would have had if we had regularly felt empathic concern for individuals affected by similar actions in the past. That is, we should defer to the judgments we have most reason to believe we would have had if we had collected the evidence provided by empathic concern in the way dictated by the habitual view.

**Section 6: Conclusion**

I have argued that although empathic concern has limits, these limits do not prevent it from ever being useful. Some of these are cases where we do not come up against these limits. For example, we do not need to worry about the closeness bias, innumeracy, burnout, or ceiling effects when the only interests at stake are our own and someone to whom we are already close. That being said, most interesting, controversial moral issues will be cases where we do brush up against the limits of empathic concern. But there are ways of overcoming these limits. For example, we can eliminate the closeness bias by using virtual reality experiences to acquaint ourselves with the suffering of others. We can overcome innumeracy by breaking up questions about actions that affect multiple individuals into smaller questions about how to treat each of those individuals. We can overcome burnout with training, or by deferring to the judgments of those who, for whatever reason, are not subject to burnout. And we can overcome ceiling effects by focusing on breaking up large misfortunes into smaller misfortunes and focusing on our reasons to prevent those small misfortunes.

In addition to specifying a set of conditions under which the limits of empathic concern do not undermine its epistemic utility, I have also discussed different ways in which we might go about collecting the evidence provided by empathic concern and incorporating it into our judgment. Each of these methods have their own advantages and disadvantages. But I have argued that the first step in properly collecting the evidence provided by empathic concern is to use the one-off view, in cases where it is feasible. When it is not feasible, or when using this method does not change our judgment, we should then turn to habitual view.

Although I have provided some abstract guidelines about how we can go about collecting the evidence provided by empathic concern in the proper way—in a way we can expect will improve moral judgment—the reader may be understandably skeptical that there are any such cases in which we can follow these guidelines. Or they allow that there are uninteresting cases in which we can follow such guidelines—such as the case discussed above where we are deciding whether to make some sacrifice to benefit some ingroup member—while suspecting empathic concern will not help us make progress on any cases of significant interest. In the second half of the dissertation, I will argue that this suspicion is not confirmed. We can use the guidelines defended in this chapter to make progress on at least two significant moral debates regarding the nature of our obligations to prevent animal suffering and fight extreme poverty.

## Chapter 3

## Empathic Concern and Eating Animals

Most of us believe that what we choose to eat is merely a matter of personal preference. Chicken or fish? Whichever you wish. Beef or pork? Put either on your fork. Vegetarian or carnivore? What're you in the mood for? Nevertheless, a small but growing contingent of people believe that what we choose to eat is in fact a matter of great moral significance. In particular, animal advocates hold that most affluent people are constantly committing egregious moral violations whenever they sit down to a meal of animal flesh. Some go so far as to argue that even consuming animal *byproducts* (e.g., milk and eggs) constitutes a grave injustice.

Animal advocates offer a number of arguments for their positions. But perhaps the most powerful style of argument—and certainly the most popular argument among non-philosophers—is the argument from species overlap. Arguments from species overlap all share the same structure. First, they point out that we would never treat members from a certain species—typically cats or dogs—in the way we currently treat farmed animals (i.e., purchasing products that lead them to be raised on factory farms), and that we in fact think it would be wrong to do so. Second, they point out that this holds true even in cases of species overlap, where the members of these species have psychological capacities that are no more complex than the psychological capacities of farmed animals (e.g., chickens, pigs). Next, they argue that there is no morally significant difference between these overlapping individuals and farmed animals such that it would be permissible to treat farmed animals one way but impermissible to treat these individuals the same way. Finally, they argue that since it would be wrong to treat these overlapping individuals the way we currently treat farmed animals and there is no morally

significant difference between these overlapping individuals and farmed animals, it must be wrong to treat farmed animals the way we currently do.

Arguments from species overlap have two types of detractors. First, some deny that it is wrong to treat companion animals with overlapping psychological capacities the same way we treat farmed animals. They think that, it would be perfectly permissible to buy food that was produced by overlapping dogs that were treated in the way farmed animals are treated, all else being equal (i.e., people not getting upset by the fact that their food comes from dogs). These are the *bullet biters*. Second, some hold that there is in fact a morally relevant difference between dogs and the farmed animals who have the same psychological capacities. They believe it would in fact be wrong to purchase food that was produced by overlapping dogs who were treated in the way farmed animals are treated, even if all else were equal. These are the *wishful thinkers*.

In this chapter, I draw on the views defended in chapters 1 and 2 to argue against both bullet biters and wishful thinkers. Against bullet biters, I argue that we have decisive reason to trust our original intuitions about the proper treatment of overlapping companion animals and to distrust our intuitions about the proper treatment of farmed animals. Our intuitions about the proper treatment of overlapping companion animals are better informed by the evidence provided by empathic concern than our intuitions about the proper treatment of farmed animals. So we should believe that it is wrong to consume standard factory-farmed products, just as it would be wrong to consume factory-farmed cat and dog products.

Against wishful thinkers, I argue that we should defer to the judgments we would have had about the permissible treatment of farmed animals if empathic concern had influenced our moral judgments in the proper way, and that if empathic concern had influenced our moral

judgments in the proper way we would believe it is wrong to consume factory-farmed animal products.

**Section 1: Against Bullet Biters**

Animal advocates argue that we would never dream of inflicting on companion animals the kinds of harms we inflict on farmed animals, even if doing so would bring us the kinds of benefits we get from inflicting harm on farmed animals (Singer 1975; Dombrowski 1997; Norcross 2004; McPherson 2014; Horta 2014). They then argue that there's no morally relevant difference between companion animals like cats and dogs and farmed animals like pigs and cows. Because there's no morally relevant difference between companion animals and farmed animals, either it's okay to treat companion animals as we currently treat farmed animals, or it isn't okay to treat farmed animals as we currently treat them. Since it's clearly not okay to treat companion animals as we currently treat farmed animals, it must not be okay to treat farmed animals as we currently treat them.

Wishful thinkers try to resist this argument by trying to find some morally relevant difference between privileged overlapping individuals and farmed animals that would make it okay to harm the latter in a way that it's not okay to harm the former. For example, they might argue that it is wrong to harm overlapping companion animals only because we care about them. Or they might argue that overlapping companion animals have morally relevant *modal* properties that farmed animals lack (Kagan 2016). I will consider wishful thinkers in the next section.

Unlike wishful thinkers, bullet biters grant that there is no morally relevant difference between farmed animals and companion animals with overlapping psychological capacities (Frey

1980; Francis & Norman 1978). So they grant that if it's wrong to treat companion animals the way we currently treat farmed animals, it's also wrong to treat farmed animals that way. But they modus tollens where animal advocates modus ponens. Rather than concluding that it's wrong to treat farmed animals the way we currently do, they conclude that it is in principle permissible to treat companion animals the way we treat farmed animals when doing so would bring us comparable benefits (Cf. Norcross 2004).

At this point, animal advocates and bullet biters reach a stalemate. Both parties agree that we must reject one of two intuitions. But they reject different intuitions. Animal advocates reject the intuition that eating animal products is permissible. Bullet biters reject the intuition that it is wrong to treat companion animals the way we treat farmed animals.

Which intuition ought we to reject? Against animal advocates, it could be argued that we should distrust our intuition about the permissible treatment of companion animals. It might be argued that we only care about these individuals more than farmed animals because we feel empathic concern, and we only feel empathic concern because we have been distracted by certain morally irrelevant features of companions animals (e.g., how cute they are, the fact that we have grown up with them). Since cuteness and these other irrelevant features do not change an individual's moral status, they might argue that this should lead us to distrust these intuitions. However, since such irrelevant influences do not influence our intuitions about the proper treatment of farmed animals, they might hold that we can continue to trust those intuitions.

If the arguments I have given about the epistemic significance of empathic concern are correct, then bullet biters have things exactly backwards. We should privilege our intuitions about overlapping companion animals precisely because our intuitions about them have been influenced by these allegedly irrelevant influences. The fact that these allegedly irrelevant factors

have influenced our judgments just shows that these judgments were formed under favorable conditions. Just as having our eyes open and having good lighting are the conditions that enable our visual experiences to provide good evidence regarding the middle-sized dry goods in our immediate vicinity, these conditions allow our empathic concern to provide good evidence regarding the strength of our reasons to not harm these individuals. It seems that real explanation for why we so strongly recoil from the thought that it would be permissible to harm an overlapping companion animal by buying factory farmed companion animal products is that we feel empathic concern for these individuals when we imagine them being harmed, and we subsequently treat it as good evidence that we have strong reason to refrain from doing what harms them. It could be that we collect this evidence in the way the synchronic one-shot view describes or the habitual view describes. For our purposes here, it does does not matter. What matters is that our intuitions about the treatment of overlapping companion animals is influenced by empathic concern felt under favorable conditions—we see the face of the target, the face is cute, we have interacted in positive ways with these individuals, etc.—and we have no reason to doubt that it provides the good evidence that we treat it as providing. So it is our intuition that it is permissible to harm farmed animals by buying standard factory-farmed animal products that we should reject, not our intuition that it would be impermissible to harm overlapping companion animals by buying factory-farmed companion animal products.

**Section 2: Against Wishful Thinkers**

The standard dialectic between animal advocates and wishful thinkers takes on a characteristic form. The wishful thinker offers some non-normative difference between

psychologically overlapping companion animals, on the one hand, and farmed animals on the other. They then argue that this their awareness of this non-normative difference justifies them in thinking they we need not treat farmed animals as well as we think we ought to treat overlapping companion animals. It is okay to harm factory-farmed animals by buying standard factory-farmed animal products, even if it would not be okay to harm overlapping companion animals by buying factory-farmed companion animal products. The animal advocate then tries to parry this response, arguing that the difference is not significant, or not significant enough to do the work the wishful thinker wants it to.

The animal advocate's strategy in this back-and-forth is that of whack-a-mole: whenever a wishful thinker offers some putatively relevant difference between overlapping individuals and farmed animals, knock it down by showing that this difference cannot be relevant in the way that would explain why it's permissible to treat these individuals in different ways. Although not a horrible strategy, deploying whack-a-mole carries with it a certain uneasiness. The uneasiness arises from the fact that there always remains the possibility that some mole has been overlooked, that some mole will appear down the road that cannot be whacked. That is, there may be some morally significant difference between overlapping individuals and farmed animals that we are aware of but cannot articulate. Fans of whack-a-mole rely on the assumption that our failure to articulate a morally significant difference between overlapping individuals and farmed animals shows that we must not be aware of one, and so we must not be justified in believing that there is one. But this assumption can be rightfully questioned. Even if we cannot articulate the difference between farmed animals and overlapping individuals, we may still be aware of one, and this awareness may justify us in thinking that we need not treat farmed animals the same way we treat overlapping individuals. If we assumed people must be able to articulate all of

the morally relevant differences that explain why they have different judgments about different cases in order to be justified in having those different judgments, most people who have never taken an ethics class would be unjustified in holding most of their moral judgments. But most people are not unjustified in holding most of their moral judgments. At least not for this reason.

In lieu of whack-a-mole, I pursue a different kind of argument against wishful thinkers. Rather than replying in a piecemeal way to each proposal the wishful thinker might offer, I offer an argument that does not depend on the questionable assumption that our failure to articulate a morally significant difference of the right kind shows we are not aware of one. I argue that there must not be a morally significant difference that we are aware of, because, all else being equal, if we had collected the evidence provided by empathic concern in the proper way, we would not think think there was such a difference. We would think that it is wrong to eat animal products. And since we have no reason to think that collecting the evidence provided by empathic concern would lead to any kind of epistemic disadvantage, we have decisive reason to defer to the judgments we would have had under these conditions.

This argument bypasses the need to assume that our failure to articulate a morally significant difference shows that there is no such difference. Why? Because if we are justified in believing there is a morally relevant difference between farmed animals and overlapping individuals, we must at some level already know what this difference is. (To deny this would be to embrace a radical and rightfully maligned form of intuitionism on which we have the capacity to directly detect moral facts—to know them without knowing the non-moral facts in virtue of which they obtain.) But *even if we are aware of some difference that explains our different intuitions about the treatment of farmed animals and overlapping individuals*, we would come to think this difference is irrelevant if we collected the evidence provided by empathic concern in

the proper way. We need not explicitly show that any putative difference is morally irrelevant in the piecemeal way the whack-a-mole strategy does. We can show that any putative difference we are already aware of is morally irrelevant by the fact that more informed versions of ourselves would deny there is a morally relevant difference.

I argued in chapter 2 that there are various ways we could collect the evidence provided by empathic concern. I argued that if we are considering a genuinely new type of action in which the only individuals whose interests are at stake are your own and those of one other individual, then the best strategy is to follow the instructions of the *synchronic one-off view*, because following the instructions of the synchronic one-off view allows us to directly consult the evidence provided by empathic concern as we form our moral judgment without taking our attention away from other relevant evidence. It is a case where we have tunnel vision, but everything worth looking at is within the tunnel. It follows from this that there are two kinds of situations where we should not rely solely on the instructions of the synchronic one-off view. The first is a situation where we are evaluating an action of a sort we have performed or observed others perform many times in the past, and where we have not followed the instructions of the habitual view (i.e., where we have not regularly felt empathic concern for all of those who were affected by the action under favorable conditions). The second is a situation where the interests of individuals other than yourself and one other individual are at stake; where there are relevant considerations that lie beyond the tunnel.

If the action we are evaluating is one where we have a long history with the action type in question and we did not follow the instructions of the habitual view, then we may first attempt to follow the instructions of the synchronic one-off view if it is a case where only our interests and the interests of one other individual are at stake. If following the instructions of the synchronic

one-off view changes our judgment about the permissibility of the act, then we can rest content. But if following the instructions of the synchronic one-off view does not change our judgment, this may only be due to the fact that we have a long history of acquainting ourselves with a biased body of evidence. It may be that it is only because we have failed to feel empathic concern with those harmed by such actions in the past that we continue believing this action to be permissible, even as we feel empathic concern for the individuals harmed in the present case. Due to this fact, we should attempt to follow the habitual view if following the synchronic one-off view does not change our judgment. This does not mean that we must go ahead and perform the action over and over again so that we may feel empathic concern by those who are harmed by it. If we do not want to perform the action—perhaps because we are in doubt whether it is permissible to do so—then we can instead *simulate* performing the action and *simulate* feeling empathic concern for those affected. But if we are going to perform the action anyway, then we should feel the empathic concern by those who are harmed by it.

I invite the reader to go ahead and try out the synchronic one-off view themselves. For example, participate in a virtual reality experience depicting the life of an animal on a factory farm, such as Animal Equality's iAnimal (available at ianimal360.com). If doing so leads you to think that eating animal products is generally wrong, you can stop reading this chapter. You have decisive reason to accept this judgment, as your judgment will be informed by the evidence provided by empathic concern, and you have no reason to think that gaining this evidence has led to any kind of epistemic disadvantage.

But if witnessing the suffering of factory-farmed animals under favorable conditions does not lead you to think eating animal factory-farmed animal products is morally impermissible, I will now provide the argument that if we had collected the evidence provided by empathic

concern in the proper way, we would think that it is impermissible to (purchase and) eat animal products. Here it is:

(1) We should endorse the judgments we have decisive reason to believe we would have about the moral permissibility of eating animal products if we had collected the evidence provided by empathic concern in the proper way and all else were equal.
(2) We have decisive reason to believe that if we had collected the evidence provided by empathic concern in the proper way and all else were equal, we would judge that eating animal products is morally impermissible.
(3) Therefore, we should endorse the judgment that eating animal products is morally impermissible.

Before discussing these premises, let me offer several clarifications about the conclusion. First, by "eating animal products" I mean eating the flesh of chickens, pigs, cows, goats, horses, and fish raised according to standard practices within industrial animal agriculture, along with eating other products produced by these animals. What unifies these animals is that they have the capacity to suffer both physically and psychologically, and they do tend to suffer in these ways throughout their lives on factory farms. While it is beyond the scope of this dissertation to defend these claims here, I assume here that factory-farmed animals do not have lives worth living. They have very few pleasures in life and many pains, such that a representative moment in the life of a factory farmed animal is dominated by suffering of some form or other. The argument I offer has no bearing on the permissibility of eating farmed animals whose lives are worth living, of which I take there to be very few.

Second, the claim that eating animal products is morally impermissible is a generic claim. It can be true even if there are exceptions, cases where eating an animal product is morally permissible. While the truth conditions of generics remain controversial, I intend this generic claim to be interpreted as one that says something about a characteristic instance of eating animal products, where a characteristic instance of eating an animal product is something like eating an egg, or a serving of meat, or a glass of milk. It is consumption on this scale that I am interested in. I am interested in consumption on this scale because this is the scale on which most consumption decisions have to be made, and which plausibly stands some chance of causing animal suffering (although the question of whether this assumption is true will be discussed at the end of the paper). I have no interest here in defending the view that eating trace amounts of animal products is morally wrong when it stands no chance of increasing demand in such a way as to increase the number of animals raised on factory farms.

I defended premise (1) in chapters 1 and 2, so I will take that premise for granted here. But why should we believe (2)? Here is an argument in favor of (2):

(4) We have decisive reason to believe that if we collected the evidence provide by empathic concern in the proper way and all else were equal, then we would have to witness the suffering of the animals harmed by our consumption of animal products under favorable conditions whenever we consumed animal products.

(5) We have decisive reason to believe that if we witnessed the suffering of the animals harmed by our consumption of animal products under favorable conditions whenever we consumed animal products and all else were equal, then we would judge that eating animal products is morally impermissible.

(2) Therefore, we have decisive reason to believe that if we collected the evidence provide by empathic concern in the proper way and all else were equal, then we would judge that eating animal products is morally impermissible.

Premise (4) tells us what it would take for the antecedent of premise (2) to be satisfied. I defended this view in chapter 2, so I will take it for granted here. But why should we accept premise (5)?

Premise (5) claims we have decisive reason to accept an empirical prediction: that if you subject people's reinforcement learning mechanisms to a constant diet of consuming animal products in the presence of graphic imagery depicting animal suffering under favorable conditions, they would think that eating animal products is morally impermissible. The second half of this prediction falls out of the Empathic Foundations Hypothesis defended in chapter 1, together with the assumption that people would feel empathic concern for the animals who suffered for their meal if they were exposed to that suffering under favorable conditions.

We cannot test the premise (5) directly. Consider what it would take to collect the evidence provided by empathic concern in the way suggested by (5). To take just one example, whenever we ate an egg, we would have to witness the 24 hours of suffering an egg-laying hen puts in on a factory farm to produce that egg. Or suppose we ate a quarter pound of chicken. The average broiler chicken lives for 42 days on a factory farm. Collecting the evidence provided by empathic concern while we ate a quarter-pound of chicken would require us to witness a two and a half day long, representative sample of a broiler chicken's life on a factory farm. Even if we could somehow manage to represent all of this suffering in a film clip that lasts no longer than

our meal, it would not be feasible for most of us to expose ourselves to such film clips every time we sat down to eat.

So if we want to know whether (5) is true, we will need less direct evidence. Fortunately, we have such evidence. The most vivid evidence of this kind is anecdotal: watch a video depicting the suffering of factory-farmed cows, pigs, and chickens. Better yet, acquaint yourself with their suffering through virtual reality, after visiting a farmed animal sanctuary and having some positive interactions with some cows, pigs, and chickens, for only then will you find yourself in the favorable conditions for collecting the highest quality empathic evidence.

If the non-psychopathic reader were to carry out these instructions, I have little doubt that they would feel empathic concern for these animals under these conditions. But we can also look to cognitive science to see whether viewing evidence of farmed animal suffering under favorable conditions triggers empathic concern. And it does. Keeping the head on a roasted pig triggers empathic concern and reduced consumption of the pig (Kunst & Hohle 2016, Study 2a; Kunst & Palacios Haugestad 2018), presenting a living animal in a meat advertisement increases empathic concern and decreases willingness to eat meat (Kunst & Hohle 2016, Study 3), and even using the terms “cow/pig” in place of “beef/pork” increases empathic concern, which in turn leads to decreased willingness to eat meat and increased willingness to eat vegetarian instead (Kunst & Hole 2016, Study 5).

Recall that we are considering whether people would feel empathic concern for farmed animals under favorable conditions, which includes perceiving the farmed animals as cute. As might be expected, the cuter farmed animals are perceived to be, the more empathic concern they elicit from participants, and the less participants are willing to eat these animals (Zickfeld et al. 2018). So we have substantial evidence now that witnessing the suffering of farm animals leads

to empathic concern for farmed animals, empathic concern strong enough to reduce willingness to eat products derived from those farmed animals.

Even so, it may be objected that even if witnessing farmed animal suffering under favorable conditions decreases willingness to eat meat in the moment, repeated exposure to farmed animal suffering under these conditions may not elicit empathic concern, and so it may not lead to sustained aversion to animal products. Alternatively, it might be argued that people are only unwilling to eat meat when witnessing farmed animal suffering because they are disgusted, or because feeling empathic concern itself is unpleasant, and so it would not lead to sustained aversion to eating animal products, an aversion to eating animal products in the absence of such imagery.

Take the first objection: that people would not continue to feel empathic concern after repeated exposure. Suppose this is right. This is irrelevant. The empathic concern we feel for particular individuals provides us with evidence about the strength of our *pro tanto* reasons to help or not harm them, reasons grounded entirely in the nature of their misfortune. Empathic concern is a *fitting* response to the misfortune of others, and its fittingness does not depend on how many others we have felt empathic concern for recently. The fact that we have been repeatedly exposed to animal suffering in the past does not change the nature of their misfortune. So we should think that the pro tanto reason we have to help the nth animal, who does not elicit empathic concern in us, is just as strong as the pro tanto reason we have to help the first animal, who does elicit empathic concern. This means that if we wanted our moral judgments to be informed by the evidence provided by empathic concern, it would not be enough for us to repeatedly witness animal suffering under favorable conditions. We would have to repeatedly witness animal suffering under favorable conditions *and feel empathic concern*. The fact that

most of us would get burnt out is neither here nor there. It is not metaphysically impossible for us to repeatedly witness animal suffering and not get burnt out. And what premise (5) really requires is for idealized versions of ourselves who wouldn't get burnt out to think eating animal products is morally impermissible.

This leads us to our second objection: why think that feeling full empathic concern time and time again would lead us to sustained aversion to eating animal products? After all, no one actually does witness animal suffering under favorable conditions whenever they consume animal products. How could we know the appetites and aversions of people who did?

Even though no humans repeatedly witness animal suffering under favorable conditions when consuming animal products, studies have been done on animals who had to witness conspecifics suffering whenever they consumed food. For example, when rats witness other rats being shocked under favorable conditions whenever they pull a lever for food, they come to avoid pulling the lever (Church 1959; Greene 1968). Likewise, Weehkin et al. (1964) and Masserman et al. (1964) find that rhesus macaques forego food when consumption is paired with shocks delivered to other macaques. Indeed, Weehkin et al. report that the majority of macaques who had themselves experienced the shocks that they witnessed others experience starved themselves for over a week when eating was associated with shocking others. One even avoided food for a full 12 days. Similar effects have been found in pigeons (Watanabe & Ono 1986),

Why think these studies on animals tells us anything about the moral judgments humans would make if they were fed an appropriate "diet" of animal products and vivid evidence of animal suffering? It might be that all these studies show is that we would lose our appetite for animal products if their consumption of those products were paired with imagery depicting the suffering of the animals who became or produced those products. But even if we did lose their

appetite for animal products, this does not mean we would think it's *wrong* to consume those animal products. After all, people lose their appetite for foods that have made them sick in the past all the time, and they don't condemn others for eating those foods. And so even if people would lose their appetite for animal products after having collected the evidence provided by empathic concern, they might just think it's "not for them", not that it's morally wrong.

There are four reasons to think this would not be the case. First, even if feeling empathic concern for animal suffering is unpleasant, it isn't always so unpleasant as to make the entire experience of eating animal products undesirable. Indeed, if it were, we would expect people to be unwilling to eat and watch television shows depicting graphic violence at the same time. But people do this all of the time without any problem whatsoever. They do not stop eating. Or if they do stop eating momentarily, they do not stop eating that particular food in the future. Once the violent imagery disappears, they have no trouble resuming their meal. This is not the case in the studies cited above—rats and rhesus macaques continue to avoid pushing the lever for food for some time after they witness the distress of a conspecific paired with pushing the lever. This suggests that when we lose our appetite for a food when it is paired with the suffering of others, this is because we have seen that our consumption of that food has undesirable consequences not just for ourselves, but for others. And since the undesirable consequences of consuming that food are not specific to us, we will likely think that *everyone* has reason to avoid consuming that food, not just us.

Second, the negative affect we feel when witnessing animals suffer is not just *any* kind of negative affect. It's a specific kind: empathic concern. And I have already provided extensive evidence that when empathic concern serves as input to reinforcement learning, the subsequent avoidance reactions tend to be moralized. People take their own aversion to performing actions

that produce empathic-concern-triggering distress as evidence that performing that action would be wrong, not just for them but for others (Cushman et al. 2012; Miller et al. 2014; Wiech et al. 2013; Miller & Cushman 2013; Sarlo et al. 2014; Patil 2015), and this aversion is based on having felt empathic concern for victims who were harmed by similar actions in the past (See chapter 1; Blair 2013; Blair 2017). So we should expect people to moralize their aversion to eating animal products, even if they do not moralize all food aversions.

Third, as Greene (2017) argues, the neural circuitry that leads people to judge it wrong to harm humans is the same kind of neural circuitry that causes animals to refrain from the kind of harmful food-seeking behaviors at issue in the studies discussed above (Blair 2007; Crockett et al. 2010; Glenn et al. 2009; LeDoux 2000; Phelps et al. 2004; Shenhav & Greene 2014, as cited in Greene 2017). If activation of this circuitry leads to moral condemnation of actions that harm animals, we have no reason to doubt that similar activation of this circuitry—the activation we would see if people's consumption of animal products were paired with vivid depictions of animal suffering—would lead to moral condemnation of actions that harm animals, actions like eating animal products.

Lastly, even though we cannot actually run a study in which people are repeatedly exposed to animal suffering under favorable conditions when they eat animal products, we can look at factors we would expect to be correlated with people's tendency to think about and feel empathic concern for animal suffering when consuming animal products. One such factor is how much people are attached to childhood pets. And as a matter of fact, people who report having been more attached to their childhood pets are more likely to be vegetarians, and this effect is fully mediated by people's empathic concern for animals (Rothgerber & Mican 2014). In other words, while we find a correlation between how attached people were to pets during their

childhood and their current avoidance of meat, this correlation disappears when controlling for empathy felt for animals. Moreover, those who report high attachment to pets during childhood but still ate meat explained their meat consumption by agreeing with statements such as "When I look at meat, I try hard not to connect it with an animal," "I do not like to think about where my meat comes from," "I try not to think about what goes on in slaughterhouses," and "When I eat meat, I try not to think about the life of the animal I am eating" (Rothgerber 2012, p. 13). This provides striking support for the view that those who feel empathic concern for farmed animals most frequently when they consider eating animal products—i.e., those who most closely approximate the proper way of collecting the evidence provided by empathic concern—do in fact avoid eating animal products.

So people who were more attached to pets during childhood and thus feel more empathy for farmed animals over the course of their life are more likely to avoid meat, if they regularly think about the suffering of the animals when they consider eating meat. But are they more likely to judge eating animal products to be *impermissible*? It could be that the prospect of eating meat merely makes them feel uncomfortable. Or perhaps they think that vegetarianism is supererogatory rather than morally required (Harman 2015). The majority of vegetarians report avoiding meat for moral reasons (Rozin et al. 1997; Beardsworth & Keil 1991; Fox & Ward 2008; Hussar & Harris 2009; Pribis et al. 2010), but one could think these moral reasons speak in favor of avoiding meat without requiring it. Indeed, children who go vegetarian for moral reasons (e.g., because eating meat involves animal suffering and premature death) do not condemn others for eating meat unless they have already committed to abstaining from meat (Hussar & Harris 2009). Unfortunately, few studies explicitly ask adult vegans and vegetarians for deontic judgments regarding the permissibility of eating animal products. But those that do

find vegans and vegetarians typically endorse claims like "There is never a good enough of reason to violate my diet" and "I must always follow my diet completely at all times." (Rothgerber 2015a; 2015b) These findings strongly suggest that those who feel empathic concern for farmed animals most frequently when they consider eating animal products—i.e., those who closely approximate the proper way of collecting the evidence provided by empathic concern—not only avoid eating meat, but think it is wrong to eat meat and other animal products. This concludes my argument for premise (5), and thus my argument for premise (2), and thus my argument for (3): we should endorse the conclusion that eating animal products is morally impermissible.

**Section 3: Objections**

At the end of chapter 1, I argued that we should trust empathic concern under a variety of "favorable" conditions, such as when we see the face of the target individual, their face is cute, and so on. And the arguments I have given against bullet biters and wishful thinkers both rely on the crucial assumption that we should trust the evidence provided by empathic concern under such conditions. If empathic concern did not deliver good evidence under these conditions, then we would no longer have reason to trust our intuitions about the proper treatment of dogs and cats over our intuitions about the proper treatment of farmed animals. Likewise, we would no longer have reason to think that our judgments about the proper treatment of farmed animals would change if we had felt empathic concern in the proper way. So one way to attack the arguments I have given in this chapter is to question the assumption that the conditions I have called favorable conditions are in fact favorable conditions.

I argued that these conditions were favorable conditions because if they were not favorable conditions we would have to think that, all else being equal, people who did not feel empathic concern under such conditions (or else felt it but did not treat it as evidence) would enjoy an epistemic advantage relative to those of us who felt and treated empathic concern in the normal way. But, I argued that people who do not feel or treat empathic concern in the normal way clearly do not enjoy an epistemic advantage, because they seem to endorse our core moral beliefs less strongly than we think they ought to. So, I concluded, the conditions I listed must be favorable conditions after all.

In reply to this argument, one could argue that the conditions I identified as the favorable conditions were in fact narrower than I let on. I claimed that empathic concern must be a good source of evidence regarding the strength of our reasons to help and not harm a target whenever we see the target's face, their face is cute, etc. I claimed that only if this were true would we be justified in holding our core moral beliefs. But to think we are justified in holding our core moral beliefs, we need not think that empathic concern provides evidence *whenever* we see a target's face, their face is cute, etc. We need only think it's a good source of evidence under these conditions *when the target is a human being*. This is because our core beliefs are beliefs about the proper treatment of *human beings*, and we only formed these beliefs because we treated empathic concern *felt for human beings* as evidence over the course of our moral development. If we allow that empathic concern is a good source of evidence under favorable conditions *when the target of the empathic concern is a human being*, then this is thereby enough to successfully explain why we are justified in holding our core beliefs. We need not also assume that empathic concern is also a good source of evidence *when the target of empathic concern is a nonhuman animal*.

How plausible is it to narrow in this way the favorable conditions under which empathic concern provides good evidence? I believe it is plausible only if in feeling empathic concern for individuals, we assume they are humans. In that case, a wishful thinker could legitimately complain that empathic concern is unreliable when the target of empathic concern is not a human, because in feeling empathic concern for nonhumans, we are anthropomorphizing. We are holding false beliefs about the nature of the target. And it is plausible to think we should not trust empathic concern that is based on false beliefs.

But how plausible is it to think we are anthropomorphizing when we feel empathic concern for nonhuman animals? I submit that this is not plausible. When we feel empathic concern for animals who are suffering, we are feeling empathic concern because they are experiencing phenomenal states whose phenomenal character is undesirable, or because they have frustrated desires. We are not anthropomorphizing when we represent animals as having phenomenal states with phenomenal character that is undesirable, or when we represent them as having frustrated desires. These are not the kinds of mental states that only humans can have. That being said, it is beyond the scope of this dissertation to argue that the nonhuman animals who we take to suffer actually do suffer. There may be some cases in which behavior we think indicates suffering does not. If it turns out that we intuitively attribute suffering to animals when we are exposed to that suffering under favorable conditions and yet we have decisive reason to believe those animals do not suffer, then this provides us with reason to distrust empathic concern directed at such animals. But again, I take our intuitive attribution of suffering to factory farmed animals to be veridical.

The wishful thinker may object that even if we have no false beliefs about the mental states of nonhuman animals, empathic concern leads us to falsely believe that those mental states

carry the same moral significance as similar mental states had by human beings. But why should we believe this? The wishful thinker is at this point flat-footedly denying that the evidence provided by empathic concern is reliable when the target of empathic concern is a nonhuman animal. This goes hardly beyond assuming the very thesis they intend to defend. Moreover, this hypothesis is strikingly ad hoc. Saying that empathic concern is a reliable source of evidence when aimed at humans but not when aimed at nonhumans is like saying our eyes are a reliable source of evidence when aimed at tables but not chairs. The reason it is implausible to think our eyes are a reliable source of evidence when aimed at tables but not chairs is that every reason we have to trust our eyes when they are aimed at tables is a reason to trust them when aimed at chairs. The reason we have to trust our eyes when they are aimed at chairs is that these experiences seem veridical, they're clear and distinct, and we have no reason to doubt that things are as they seem. Likewise, empathic concern seems to be telling us something about how we should treat animals, just as it seems to telling us something about how we should treat humans. Every reason we have for trusting empathic concern when it's pointed at humans are reasons to trust it when it's pointed at nonhumans. From the perspective of someone experiencing empathic concern, it seems to provide evidence that is just as compelling in both cases, just as the evidence of our eyes seems just as compelling in the case of tables and chairs. We have no non-question-begging reason to doubt that empathic concern is providing the evidence it seems to when it is aimed at animals.

I have been operating on the assumption that our core moral beliefs are beliefs regarding the proper treatment of *human beings*. But in so doing, I was being overly concessive. Many of our core moral beliefs pertain to the proper treatment of nonhuman animals like cats and dogs. It is uncontroversial that it is wrong to light a cat on fire for fun. This is something we take

ourselves to be justified in believing, and we have no reason to doubt we are justified in believing it. But like our other core moral beliefs, we should think we only have this core belief because we have felt empathic concern for cats in the past. Since we take ourselves to be justified in holding this core belief about cats, we must thereby think that empathic concern provides high quality evidence regarding the strength of our reasons to help and not harm cats. But once we've allowed empathic concern to give us high quality evidence regarding how we should treat companion animals like cats and dogs, why deny that it gives us high quality evidence regarding how we should treat farmed animals like pigs and chickens? If it seemed ad hoc to limit our trust in empathic concern to cases where it's pointed at humans rather than nonhumans, it seems even more ad hoc to limit our trust in empathic concern to companion animals and not farmed animals.

Let us now move onto a second objection. It might be argued that premise (5) is false—we would not come to believe that eating animal products is morally impermissible if we repeatedly felt empathic concern for the animals who were harmed by our consumption of animal products—because *no animals are harmed by our consumption of animal products*. According to this objection, an individual's consumption behavior does not make a difference to whether any animals suffer, because their contribution to demand is too small to affect farmers' decisions about how many animals to raise (Budolfson in press; Nefsky 2011; cf. Sinnott-Armstrong 2005). Given the magnitude of decisions that have to be made about how many animals to raise on factory farms, decision makers will never raise fewer animals simply because you left a serving of meat that you would have otherwise eaten for dinner on the grocery store shelf. They likely won't know you abstained from meat, and even if they did it wouldn't change

how many animals they raise. They expect there to be small fluctuations in demand, and so small decreases in demand will not change how much they anticipate being able to sell in the future.

A number of philosophers have argued that, although it is unlikely that our particular purchasing decisions will make a difference to how much animal suffering occurs, there is *some* chance they will, and if they do make a difference, they will make a big enough difference that the expected utility of the purchase will be equivalent to however much suffering was required to produce the product in question. For example, Kagan (2011) argues that there is a 1/n chance that your purchase of a chicken will trigger a decision on the part of the farmer to raise an additional n chickens. If so, then your purchase of a chicken is equivalent to causing a lifetime of suffering to a chicken on a factory farm.

I do not want to wade into this debate here. I will simply offer some conditional conclusions. First, if the expected effect of consuming animal products is equivalent to causing an animal to experience the amount of suffering required to produce that product, and if our reason to avoid performing an action with this expected effect is just as strong as our reason to avoid causing an animal to experience that amount of suffering, then premise (5) is true: there would be an animal (perhaps indeterminate) who is harmed by our consumption of animal products, and so we would feel the empathic concern needed to kickstart the learning process that would result in our moral judgment that eating animal products is typically wrong.

Second, if in fact the expected outcome of consuming animal products is not equivalent to the suffering required to produce the product, then we can still learn something from the fact that we would judge it impermissible to eat animal products if our consumption of those products caused the suffering that was required to produce it. What we learn is that our reasons to prevent animal suffering are much stronger than our reasons to ensure that people get to enjoy the taste

of animal products instead of vegan foods. So those who are in positions of power—those who *can* (non-coercively) influence how many animals raised on factory farms—have strong reason to reduce the consumption of animal products in the aggregate. Assuming their reasons to allow people to consume animal products are no greater than people's own reasons to consume animal products, they can see that the strength of their reason to prevent farmed animals from suffering is equivalent to the aggregate strength of the reasons people *would* have to avoid causing the suffering required to produce the animal products they consume, if their consumption actually caused that suffering. The aggregate strength of such reasons will be quite strong, indeed.

Third, there are a number of plausible views on which we have reason to perform actions, even when collective action problems prevent those actions from having the effects they would seem to have in isolation. For example, some believe we not only have reasons to avoid *causing* others to suffer, but we also have reason to avoid *benefiting* from the suffering of others. Others believe we have reason not only to avoid *causing* suffering, but to avoid *helping to bring about* or *jointly causing* suffering (Nefsky 2017; Harman 2015). Suppose some such story can be told that explains why we have reason to avoid "contributing" to bad outcomes in cases where our contribution nevertheless makes no difference to whether those outcomes obtain. If the strength of these reasons is roughly equivalent to the strength of the reasons we would have to avoid producing our "fair share" of the bad outcome—as most people who offer such accounts believe—then the process I describe for how we should go about collecting the evidence provided by empathic concern can still be expected to yield correct verdicts about the permissibility of eating animal products. It will yield the verdict that it is typically impermissible to eat animal products. And as premise (1) implies, we have decisive reason to accept this verdict.

Lastly, one might argue that if I am right, then those who are most intimately acquainted with the suffering of farmed animals—slaughterhouse workers and factory farmers—would be least likely to think it's morally permissible to eat animal products. But this seems to be the opposite of what we find. Despite being repeatedly exposed to the suffering of farmed animals in the most vivid way possible, they think it is okay to eat animal products.

Although farm workers are intimately acquainted with the suffering of farm animals, as a whole they are not in favorable conditions for feeling the kind of empathic concern that constitutes high quality evidence regarding the strength of our reasons. This is for several reasons. First, it is overwhelmingly plausible to believe that farm workers often have false beliefs about the mental states of farmed animals. They often falsely believe they are unintelligent automata who may not even be capable of suffering. Second, they often have to stifle their empathic concern in order to continue performing their job well. The same is true of doctors (Cheng et al. 2007; Decety et al. 2010). Third, although they find themselves in some favorable conditions—seeing the animal's faces, hearing their screams, representing them to be nearby—they are not in one of the most important favorable conditions: having had positive interactions with the animals. Most of the interactions that farm workers have with animals are highly antagonistic. Farm animals do not want to be handled by farm workers in the way that farm workers handle them. All of these considerations give us reason to discount the intuitions of factory farm workers when it comes to the moral permissibility of eating animal products. Although they are vividly acquainted with the suffering of factory-farmed animals, they do not witness this suffering under favorable conditions.

**Section 4: Conclusion**

I have argued in this chapter that we have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way, we would come to believe that eating factory-farmed animal products is typically morally impermissible. I argued further that since collecting the evidence provided by empathic concern in the proper way in these cases does not introduce any apparent epistemic disadvantage, we have decisive reason to endorse the judgments we would have under these conditions. We should think that eating factory-farmed animal products is typically impermissible.

# Chapter 4

## Empathic Concern and the Demands of Morality

On June 2nd, 2017, five bystanders stood by and watched as Jamel Dunn drowned in a pond in Cocoa, Florida. The bystanders were the only ones to witness this man drowning, and they did nothing at all to help him. They didn't dive in and try to pull him to safety. They didn't throw him a rope. They didn't even dial 911. (CBS/AP 2017)

Predictably, the world united in outrage at the gross indifference displayed by these individuals. Not only did people think that what these bystanders had done was wrong. They thought it was depraved. The bystanders' failure to help in this case was particularly egregious, since by all accounts they had nothing at all to lose in helping the drowning man. Calling 911 would have been completely costless. They would have had to sacrifice nothing at all to save his life.

But imagine helping the man would have required them to make some sacrifice. Perhaps if they had stayed and tried to help the man, they would have missed the season premiere of their favorite television show. Would this have made their behavior any more acceptable? Not at all. If anything, their offering this as an excuse for their failure would make them seem even more callous. But let's imagine helping the man would have come at a greater cost to them. Perhaps sticking around to help the man would have required them to miss a flight on an airline that has a strict policy against issuing refunds or rescheduling missed flights. Or perhaps sticking around to help the man would have prevented them from selling some stocks before their price plummeted, resulting in a loss of $10,000, which would mean having to drive a Ford Focus for the rest of the year instead of buying that new BMW they wanted. Perhaps their having to make this much of a

sacrifice would go some small way toward helping us understand how these bystanders could have failed to help, but it would do nothing to shake our conviction that their failure to help was a grave injustice.

If you're reading this dissertation, there's a good chance you make more money than you need to lead a pleasant life. What do you do with that extra income? If you're like most people, you probably spend it on some Very Nice Things. You take expensive vacations halfway across the world. You drive a luxury car rather than an economy car. You live in a home with more square feet and amenities than you need to live comfortably. You send your kid to a private university rather than the perfectly adequate state-run college in town.

If we wanted to, we could take the money we spend on Very Nice Things and use it to make the world a better place. If we gave it to the right charities, this money could do an enormous amount of good. It could save lives. It could prevent disease and disability. It could relieve untold amounts of suffering. When some people learn this, they decide to live more frugally and give the money they save to charity. But most of us continue to spend lavishly. And few of us think that our doing so is wrong. But is it? If so, how much are we required to give to charity?

In this chapter, I aim to make a contribution to the debate on this question. I will argue that we are required to give much more money to charity than most of us currently think. I will argue for this claim by relying on the account of empathic concern's epistemic significance defended in chapters 1 and 2. I will argue that if were to collect the evidence provided by empathic concern in the proper way, we would judge that we are obligated to give much more money to charity than we currently think we are. I will argue further that we have decisive reason

to defer to the judgments we should make under these conditions. So we should judge that we are obligated to give much more to charity than we currently think.

In section 1, I briefly review Peter Singer's argument for the claim that we are required to give substantial amounts of money to charity. In section 2, I briefly discuss various reasons detractors have had for resisting Singer's argument. In section 3, I argue that if we collected the evidence provided by empathic concern in the proper way, we would judge that we are obligated to give much more money to charity than we currently think. I contend that this argument avoids the objections that detractors rightly offer in response to Singer's argument. In section 4, I consider two objections to the view that we should defer to the judgments we would make if we collected the evidence provided by empathic concern in the proper way. Section 5 concludes.

**Section 1: The Arguments for Demandingness**

The case of Jamel Dunn is a real-life version of a hypothetical case first considered by Peter Singer in his famous 1972 article "Famine, Affluence, and Morality" and again in his 2009 book *The Life You Can Save*:

> On your way to work, you pass a small pond. On hot days, children sometimes play in the pond, which is only about knee-deep. The weather's cool today, though, and the hour is early, so you are surprised to see a child splashing about in the pond. As you get closer, you see that it is a very young child, just a toddler, who is flailing about, unable to stay upright or walk out of the pond. You look for the parents or babysitter, but there is no one else around. The child is unable to keep his head above the water for more than a few

> seconds at a time. If you don't wade in and pull him out, he seems likely to drown. Wading in is easy and safe, but you will ruin the new shoes you bought only a few days ago, and get your suit wet and muddy. By the time you hand the child over to someone responsible for him, and change your clothes, you'll be late for work. What should you do? (Singer 2009, p. 3)

Singer observes that most morally competent people would strongly believe that you ought to save the child in such a case, even though doing so would require you to sacrifice several hundred dollars worth of clothing. He assumes that this is the right response. He parlays this assumption into an argument for an independently plausible normative principle: "If it is in your power to present something bad from happening, without sacrificing anything nearly as important, it is wrong not to do so." (Singer 2009, p. 15; Cf. Singer 1972, p. 231) Singer then offers a body of evidence that strongly suggests that affluent individuals can prevent bad things from happening to those in extreme poverty—bad things like malnutrition, lack of shelter, and disease—by donating to effective charities working overseas. From this Singer concludes that we ought to give money to effective charities working to alleviate suffering associated with extreme poverty.

How much money? In "Famine, Affluence, and Morality," Singer writes: "...I and everyone else in similar circumstances ought to give as much as possible, that is, at least up to the point at which by giving more one would begin to cause suffering for oneself and one's dependents—perhaps even beyond this point to the point of marginal utility, at which by giving more one would cause oneself and one's dependents as much suffering as one would prevent in Bengal." (pp. 233-234) He takes this to follow from the fact that if we stopped giving at any

earlier point, we would be failing to prevent something bad from happening, even though preventing that bad thing from happening would not require us to sacrifice anything nearly as significant. In *The Life You Can Save*, he makes this more vivid:

> ...while the cost of saving one child's life by a donation to an aid organization may not be greater, after you have donated that sum, there remain more children in need of saving, each one of whom can be saved at a relatively small additional cost. Suppose you have just sent $200 to an agency that can, for that amount, save the life of a child in a developing country who would otherwise have died. You've done something really good, and all it has cost you is the price of some new clothes you didn't really need anyway. But don't celebrate your good deed by opening a bottle of champagne, or even going to a movie. The cost of that bottle or movie, added to what you could save by cutting down on a few other extravagances, would save the life of another child… So you must keep cutting back on unnecessary spending, and donating what you save, until you have reduced yourself to the point where if you give any more, you will be sacrificing something nearly as important as a child's life—like giving so much that you can no longer afford to give your children an adequate education (Singer 2009, p. 18).

As it happens, it takes much more than $200 to save the life of a child in a developing country who would die due to famine or disease. Current estimates are closer to $5,000 (Givewell 2017). Nevertheless, affluent individuals still regularly spend these amounts of money on items the sacrifice of which would be trivial compared to the benefit to a child of not dying: sacrifices like living in a smaller apartment, buying an economy car rather than a luxury car, taking fewer

vacations, having a somewhat less comfortable retirement, and sending one's children to public schools rather than private schools. Indeed, Peter Unger offers a similar argument in which the costs seem this high and yet it nevertheless seems quite plausible that the sacrifice ought to be made to save someone's life:

> *Bob's Bugatti*. On a rural road near the garage where it's securely kept, Bob's gone for a careful drive in his Bugatti. At a certain point, he spies a shiny object. To inspect it, Bob parks his car in the only place from where, directly, he can proceed on foot for a close encounter, a parking place that's just tend yards beyond the end of a certain trolley track. As it develops, when Bob walks over to the shiny object, he finds it's a switch that can be set in two ways. And, as Bob observes, there's a trolley up the line that's barreling toward the switch's fork. As the shiny switch is set, the trolley will go down the fork's opposite side, not the branch leading to a spot near Bob's Bugatti. But, as Bob sees, on that side there's a young child trapped on the track. As he knows, Bob has two options: If he does nothing about the situation, the child will be killed, but he'll enjoy a comfortable retirement. If he changes the switch's setting, his second option, then, while nobody's killed, after rolling down the vacant branch and beyond that track's end, the trolley will totally destroy Bob's uninsurable Bugatti, wiping out his entire retirement fund. Bob chooses the first option and, while the child is killed, he has a comfortable retirement. (Unger 1996, p. 136)

Quite plausibly, Bob acts wrongly here. And it is hard to spot a morally relevant difference between Bob's failure to sacrifice his Bugatti and our failure to sacrifice the money we would

spend on a comfortable retirement to save the lives of not just one, but many individuals in the developing world.

Singer's and Unger's arguments target sacrifices that would save lives. Whether these arguments support giving to save lives in the real world depends in part on how costly it is to save lives in the real world. But whatever the costs of saving lives, smaller amounts of money can still achieve large benefits short of saving one's life, benefits that are still substantially greater in moral importance than the benefits affluent individuals would get by keeping the money that could bestow those greater benefits. It is important to know how much we ought to sacrifice to bestow these non-life-saving benefits as well as how much we ought to sacrifice to bestow the benefit of avoiding premature death. So Singer's and Unger's arguments would be important to consider, even if it turned out we could not give enough to save lives without making significant sacrifices.

At the end of the day, Singer's argument is of interest regardless of how we should translate its conclusions into the real-world. It may be that there are causes even more important than the ones Singer and Unger target. What is most important about this line of argument is that it shows us something about how demanding morality could be *in principle*, how much it could require to sacrifice of ourselves to help strangers. Although adding several plausible empirical assumptions to Singer's argument can show that morality is demanding in practice, and demanding in particular ways, I consider my contribution to this debate to primarily be to the question of how demanding morality can be in principle.

**Section 2: Detractors**

In this chapter, I have no plans to respond to detractors who object to the empirical assumptions behind Singer and Unger's arguments. As just stated, I am interested here primarily in the question of how much we could be obligated to give in principle. (Nevertheless, I do believe the empirical assumptions are plausible, so I believe we do learn facts about how we ought to act in the real world. I will not defend these assumptions here.)

The kinds of objections to these arguments that I want to focus on here are objections to the principle that Singer uses the case of the Drowning Child to motivate: "If it is in your power to present something bad from happening, without sacrificing anything nearly as important, it is wrong not to do so." Some object to this principle because they independently believe that we lack obligations to help distant strangers; we have obligations only to avoid harming them (Narveson 2003)

Others allow that we have obligations to help distant strangers, but say that our obligations are limited in ways that Singer assumes they are not (e.g., Kamm 1999; Schmidtz 2000; Miller 2004; Gert 2005; James 2007; Timmerman 2015; see also Pogge 2008, p. 133, who agrees that the considerations adduced by Singer fail to ground exceptionally strong obligations to aid distant strangers, but nevertheless thinks we have exceptionally strong obligations to aid distant strangers arising from our having harmed them). Even when supplemented by an account of *why* our obligations are limited, however, objections to Singer's normative principle are typically based on just how unintuitive it is to think we ought to impoverish ourselves to help those we have not harmed, even if by doing so we bestow significant benefits on others.

Of all the objections to Singer's argument for his normative principle, Timmerman offers the most direct attack. Rather than noting how unintuitive its implications are, Timmerman argues that the normative premise that Singer offers is not the best explanation of our intuition about Singer's *Drowning Child* case. Put another way, he argues that there is a morally relevant difference between Singer's *Drowning Child* and the kind of case most of us face when we consider whether to impoverish ourselves in the name of charity. He offers what he takes to be a more analogous case:

> *Drowning Children*: Unlucky Lisa gets a call from her 24-hr bank telling her that hackers have accessed her account and are taking $200 out of it every 5 min until lisa shows up in person to put a hold on her account. Due to some legal loophole, the bank is not required to reimburse Lisa for any of the money she may lose nor will they. In fact, if her account is overdrawn, the bank will seize as much of her assets as is needed to pay the debt created by the hackers.
>
> Fortunately, for Lisa, the bank is just across the street from her work and she can get there in fewer than 5 min. She was even about to walk to the bank as part of her daily routine. On her way, Lisa notices a vast space of land covered with hundreds of newly formed shallow ponds, each of which contains a small child who will drown unless someone pulls them to safety. Lisa knows that for each child she rescues, an extra child will live who would have otherwise died. Now, it would take Lisa approximately 5 min to pull each child to safety and, in what can only be the most horrifically surreal day of her life, Lisa has to decide how many children to rescue before entering the bank. Once

> she enters the bank, all the children who have not yet been rescued will drown. (Timmerman 2015; pp. 208-209)

According to Timmerman, it is intuitively obvious that Lisa need not impoverish herself:

> I propose that it's a viable option that morality permits Lisa to, *at least* on 1 day over the course of her entire life, stop the hackers in time to enjoy some good that is not nearly as important as a child's life. Maybe Lisa wants to experience theatre one last time before she spends the remainder of her days pulling children from shallow ponds and stopping hackers. Given the totality of the sacrifice Lisa is making, morality intuitively permits Lisa to indulge in theatre *at least* one time in, let's say, the remaining 80 years of her life. In fact, commonsense morality should permit Lisa to indulge in these comparably morally insignificant goods a non-trivial number of times, though a single instance is all that is required to demonstrate that premiss two is false and, consequently, Singer's argument is unsound. (2015, p. 210)

Since it is intuitively obvious that Lisa need not impoverish herself—or at least, that she may forego making a sacrifice demanded by Singer's principle on at least one occasion—Singer's principle is false. Timmerman argues that it is in virtue of Lisa's past and future sacrifices that she is allowed to make at least this one exception, and that it is in virtue of having such a pattern of sacrifice that she is permitted to violate Singer's principle, even if we cannot violate Singer's principle in his *Drowning Child* case, in which we implicitly assume that the sacrifice we make in saving the drowning child is one-off.

Although Timmerman officially wants to show that we can permissibly violate Singer's principle on at least one occasion, it's clear how this kind of argument can be made to show that we are not required to make nearly as much of a sacrifice as Singer's principle says we ought to. We can infer that however much time and money Lisa is entitled to take away from saving drowning children is equivalent to the amount of time and money we are entitled to spend on ourselves rather than giving to effective charities. Indeed, Timmerman argues that "For what it's worth, I am inclined to hold that Lisa is obligated to rescue a great many children, though *significantly* fewer than is required by [Singer's principle]." (p. 210) At the end of the day, the argument I give in the next section is compatible with Timmerman's own intuition that we may be able to sacrifice significantly less than Singer's principle requires. But I will argue that empathic concern helps us see that we cannot rely on intuitions of the sort Timmerman relies on to show that our obligations to give are weaker than Singer believes. To discover how strong our obligations to give are, we must exploit the evidence provided by empathic concern.

**Section 3: The Argument from Empathic Concern**

My reply to detractors is straightforward. I will argue that if we collected the evidence provided by empathic concern in the proper way, we would judge that we ought to help many more individuals than detractors believe, perhaps as much as Singer's argument implies. Moreover, I argue that we ought to defer to the judgments we would have if we collected the evidence provided by empathic concern in this way. We have decisive reasons to believe that the judgments we would issue under these conditions are more likely to be correct than the judgments we have now, judgments that are not adequately informed by the evidence provided

by empathic concern. Since the detractors I am targeting object to Singer's argument based on their intuitive judgment that, for whatever reason, morality is not a demanding as Singer thinks, their arguments have no force against my view, which provides a principled explanation of why their judgments are not to be trusted over the judgments we would have if we used empathic concern in the way we ought to.

In chapter 2, I offered several ways we might go about collecting the evidence provided by empathic concern. I noted at the time that these different methods had both advantages and disadvantages. In the present case, I submit that the proper way to collect the evidence provided by empathic concern would be to follow the instructions of the *synchronic, one-shot view*. The main disadvantage of the synchronic, one-shot view—that it does not counteract the pernicious influence of our having collected empathic evidence in a biased way throughout our moral development—seems less important here than it did in the case of eating animal products, since it does not seem that the fact that we have collected empathic evidence in a biased way throughout our moral development exerts too strong an influence on our judgments about the kinds of sacrifices required by Singer's Principle. Whereas applying the synchronic, one-shot view for the first time while eating an animal product may not lead us to think that eating that animal product is wrong, it does seem that applying the synchronic, one-shot view for the first time in the domain of charitable giving (e.g., in Singer's original *Drowning Child* case) does get us to think that it would be wrong to spend frivolously on ourselves. So we should use the synchronic, one-shot view here, as it seems that its main disadvantage is not a disadvantage in this case, and yet it does give us an advantage that following the habitual view would not—being able to draw on the empathic evidence at the very time we make a judgment about the moral proposition that it is evidence for.

As I noted when introducing the view, however, the synchronic, one-shot view works best in cases where the interests of only one individual other than the moral judge and moral agent are at stake. When considering whether we should give enough money to save many individuals from premature death, which is the question we are focused on here, we cannot simply follow the synchronic, one-shot view while we are issuing our judgment about how much money we ought to give in total. Instead, we should follow the strategy I mentioned for using the synchronic, one-shot view in such cases: break up the action we are evaluating (e.g., giving away all of the money we are saving for retirement), into a series of smaller actions that affect only one individual and require a proportionally smaller sacrifice. We should then aggregate the total amount we have decisive reason to think we ought to sacrifice across all of these cases to figure out how much we ought to donate now, in a one-shot interaction, such as writing a check to charity.

Timmerman's iterated pond case, *Drowning Children*, provides an ideal test bed for using the synchronic, one-shot view in this way. To apply the synchronic, one-shot view to Lisa's case, we must imagine being Lisa (or at least watching Lisa) as she confronts each of the children drowning in each of the ponds she passes on her way to the bank. We must add that when she confronts the suffering of the drowning children, she does so under favorable conditions. She must see each of their face, hear their screams for help, and have had significant positive interactions with them in the past, the kinds of minimal positive interactions that lead people to care more for one another without generating any special obligations such individuals might have to help one another. As Lisa confronts each child under these conditions, she must ask herself whether she is obligated to save that child. If what I have argued in earlier chapters is correct, the empathic concern that Lisa will undoubtedly feel for each child will influence her

judgment in such a way as to improve it. Since the only interests at stake in this case are the interests of the drowning child and Lisa herself, it seems that the "tunnel vision" produced by feeling empathic concern in this case will not be epistemically pernicious, since everything relevant to the question of how much she ought to give—the nature of her sacrifice and the nature of the benefit to the child she could help—are within the tunnel she gazes into.

I submit that if those of us who are affluent, non-psychopathic individuals put ourselves in Lisa's position, feeling empathic concern for each individual and judging whether we ought to help that individual *while* feeling empathic concern for them, we would end up judging for many children that we ought to save them. To see the initial plausibility of this claim, simply imagine being in Lisa's position and doing this. I submit that you will imagine going through a number of iterations of feeling this process of judging whether to save a child at the cost of $200 (or even $5,000 or what have you) while feeling empathic concern under favorable conditions before you stopped judging that you ought to save the child in front of you.

Furthermore, I submit that however many children you end up judging you ought to save radically diverges from the number you actually *would* judge you ought to save if you actually *were* feeling empathic concern for these individuals under favorable conditions; imagining being under favorable conditions is not as powerful as actually being under favorable conditions. And given the arguments I provided earlier in this dissertation, the empathic evidence we gather while under favorable conditions is certainly higher quality evidence than the empathic evidence we gather when merely imagining being in favorable conditions.

In any case, we seem to have decisive reason to believe that if we put ourselves in Lisa's position, the total number of children we would judge we ought to save would be much higher than the number we currently believe we ought to save. If we follow the method argued for

earlier, we should infer that the total number of children we would judge we ought to save across these iterated cases is equivalent to the number we ought to save now. So we should infer that the total number of children we ought to save now by giving to charity is much higher than we currently think, since most of us currently think we ought to save none, or just a few.

Let me make the argument just provided maximally explicit.

(1) When it comes to the number of children we ought to save, we should endorse the judgment we have decisive reason to believe we would have if we collected the evidence provided by empathic concern in the proper way and all else were equal.
(2) We have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way and all else were equal, we would judge that we ought to save many more children than we currently believe.
(3) Therefore, we should save many more children than we currently think.

The first premise is unobjectionable. It is just the claim that we should endorse the judgments we have decisive reason to believe we would have if we had the epistemic advantage of having more relevant evidence and no outweighing epistemic disadvantages. It is the second premise that is controversial. Here is the argument for the second premise:

(4) We have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way and all else were equal, we would ask whether we should save one child (and then a second child, and then a third child...) under favorable

conditions, and then infer that the cumulative number of children we end up judging we ought to save across these cases is the number of children we ought to save right now.

(5) We have decisive reason to believe that if we asked whether we should save one child (and then a second child, and then a third child…) under favorable conditions, and inferred that the cumulative number of children we end up judging we ought to save across these cases is the number of children we ought to save right now, we would believe that we ought to save many more children right now than we currently think.

(2) Therefore, we have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way and all else were equal, we would believe we ought to save many more children right now than we currently think.

I have already given what I take to be strong initial support for (4) and (5). In the next section, I will consider objections to (4) and (5) and in the process of replying to these objections provide even stronger support for these premises.

**Section 4: Objections**

Premise (4) provides an account of what it would be to properly collect the empathic evidence bearing on the question of how much we should give to charity. It also claims that following the instructions of this account would not introduce any epistemic disadvantages that would undermine the epistemic advantage of gaining this evidence. Both objections to premise (4) question whether this ceteris paribus clause holds.

The first objection to premise (4) claims that the account on offer of how to collect the evidence provided by empathic concern is incorrect. According to this objection, we should not assume that the number of children we end up judging we ought to save across all of these cases is equivalent to the number we ought to save now, specifically because there are morally relevant differences between making a series of choices about whether to save each child when we face the choice to help them and only them at some cost, and the choice of whether to save them all at once. On this view, the sum of sacrifices we make to save each child when do so in a serial fashion is not equivalent to the sacrifice we would make if we saved all of the children in one fell swoop, even if the total amount of money we end up giving is equal. Why would this be? Because it's true that (e.g.) $200 may not make any real difference to how well our life goes. So if we judged for each of 100 children that we ought to sacrifice $200 in order to save their lives, each sacrifice of $200 would harm us almost not at all, but it would benefit the recipients a great deal. Nevertheless, giving away $20,000 all at once would harm us in a significant way, a way significant enough (so the objection goes) to make it rational to refrain from giving it all away, even to save many lives.

This is a standard kind of practical sorites case. To compare, consider Warren Quinn's "Puzzle of the Self-Torturer":

> Suppose there is a medical device that enables doctors to apply electric current to the body in increments so tiny that the patient cannot feel them. The device has 1001 settings: 0 (off) and 1 … 1000. Suppose someone (call him the self-torturer) agrees to have the device, in some conveniently portable form, attached to him in return for the following conditions: The device is initially set at 0. At the start of each week he is

> allowed a period of free experimentation in which he may try out and compare different settings, after which the dial is returned to its previous position. At any other time, he has only two options—to stay put or to advance the dial one setting. But he may advance only one step each week, and he may *never* retreat. *At each advance he gets $10,000*.
>
> Since the self-torturer cannot feel any difference in comfort between adjacent settings, he appears to have a clear and repeatable reason to increase the voltage each week. The trouble is that there *are* noticeable differences in comfort between settings that are sufficiently far apart. Indeed, if he keeps advancing, he can see that he will eventually reach settings that will be so painful that he would then gladly relinquish his fortune and return to 0. (Quinn 1990, p. 79)

Intuitively, it would be wrong to infer that since it is rational for the self-torturer to increase the voltage each week until he gets to setting 1000, it would be rational for him to go ahead and boost the voltage to setting 1000 right now, if he could do so in exchange for the same amount of money.

I submit that Lisa's situation is not relevantly like the situation of the self-torturer. The costs Lisa has to endure to save a life—whether we assume that cost is $200 or $5000—are significant. They are not as significant as the benefit of avoiding premature death, but they are significant all the same. In any case, they are not imperceptible. So we need not think that she is falling into a trap quite like that of the self-torturer.

Still, one might hold that the cumulative reason Lisa has to make her $200 sacrifice on each occasion is smaller than the reason she has to avoid giving $20,000. The fact that this

reason is sufficient to obligate her to give $200 on each occasion does not show that she is obligated to give $20,000 on one occasion.

I recognize that this is a real possibility. So this is some worry to think that there is some possible disadvantage to collecting empathic evidence in this way. But the real question is whether this epistemic disadvantage is larger than the epistemic disadvantage of proceeding without the evidence provided by empathic concern. And I submit that the epistemic disadvantage of proceeding without the evidence provided by empathic concern is quite large, indeed.

But even if the epistemic disadvantage of proceeding without the evidence provided by empathic concern is not very large, it may turn out that considering Lisa's case shows we ought to give lots of money to charity. It just shows that we ought to give lots of money to charity in small increments over time. After all, we all face the choice of whether to give small amounts every day. Even if there's no day on which it is rational to give $20,000 all at once, because doing so would require us to sacrifice something of great significance, it will continue to be rational for us to delve out $200 on each occasion that we can. And so we will still end up giving away lots of money and saving many more individuals than we initially thought we should. It will just take us longer.

The only way out of this conclusion would be to argue that the foreseeable cost of giving away $20,000 would make it rational for us to put in place measures to prevent us from considering whether to give $200 on future occasions. I will simply bite the bullet here and say that it might. But there are no practical measures we could use to do this. So it is a non-issue, as far as the practical upshot of these arguments goes.

Let me know consider a second objection to premise (4). The second objection to premise (4) is that after following this procedure of collecting empathic evidence for some time, our judgment will be perniciously influenced by a bias we will develop towards giving. More specifically, consider that in following this procedure, we will be in effect following not only the synchronic, one-off view many times in a row, but we will be following the habitual view. We will be collecting empathic evidence for a series of actions (earlier cases of giving) that are quite like later actions we will be evaluating (later cases of giving). And as discussed in chapter 2, it may be that when we collect empathic evidence in the way the habitual view instructs, we develop a sort of *habit* of treating future actions like previous actions that they are similar to. And giving $200 to the 100th child is in certain respects quite similar to giving $200 to the first few children, so we will almost certainly apply our habit of thinking we are required to give $200 that we developed on the basis of early cases to later cases of giving $200, where we have already given quite a lot. Nevertheless—and here is the heart of the objection—giving $200 when confronted with the hundredth child (e.g.) differs in morally relevant respects from giving $200 to earlier children. Due to the diminishing marginal utility of money and (plausibly) the greater normative significance of bestowing benefits to those who are worse off, we might not be obligated to give $200 when we have already given $19,800. But the habit we develop on the basis of giving in earlier cases will ignore this difference. It will treat the 100th case as if it were one where we have not made ourselves worse off and one where $200 benefits us less than it actually does (because in the 100th case we would spend it on food, whereas in earlier cases we might just spend it on unnecessary clothing).

The force of this objection depends on just how quickly the marginal utility of money declines, as well as how quickly the marginal goodness of a benefit declines according to the

wellbeing of those it goes to. I submit that for many affluent individuals, they can give many thousands of dollars before the marginal utility of the money they give away begins to increase significantly. Likewise, I submit that many affluent individuals could give quite a bit away and still remain affluent. So to the extent the present objection goes through, it only requires us to stop relying on this procedure after having given away substantial sums of money, sums of money greater than many of us already take ourselves to be required to give away.

Secondly, I do not think that we could easily develop a habit of giving in the way the objection supposes. Although we would be giving many times over in the circumstances imagined, we would not be doing it so often as we perform actions that we normally habitualize, such as (to think back to the previous chapter) eating animal products. But even if we could develop such a habit, it still remains far from obvious that having this habit influence our judgments in each case is as disadvantageous as losing the evidence provided by empathic concern. Given how important empathic concern is for helping us see obligations where we otherwise not (consider psychopaths once again), the likelihood that we would be overlooking reasons we have to give if we ignored the empathic evidence seems likely to be much more damaging than the influence of some possible habit we would give. In any case, the habit in question would lead us to think we are required to give more than we think we are, but it would not lead us to think that something impermissible is permissible. It is not plausible to think that it is *impermissible* for people to impoverish themselves by being highly charitable. On the other hand, ignoring the empathic evidence provided by empathic concern may lead us to to think that something that is impermissible (i.e., not giving away money to charity) is permissible. So it seems moral risk here is asymmetrical. It seems more important to avoid mistakenly overlooking an obligation than it is to mistakenly reify what is supererogatory into something that is required.

In light of this asymmetrical moral risk, I believe we have even more reason to prefer collecting the evidence provided by empathic concern than avoiding the allegedly pernicious habit of thinking that we ought to give even in cases where we needn't.

Let me move on now to challenges to premise (5). Premise (5) says that if we followed this procedure of putting ourselves in favorable conditions with respect to each child we could save, and asking whether to save them while in these favorable conditions, we would end up judging that we ought to save many of these children, many more than we currently think we're required to save. One way to doubt this is to doubt whether putting ourselves in these favorable conditions would in fact lead us to think we ought to help these children. The biggest reason to doubt this is the possibility of *burnout*. We may simply find ourselves no longer capable of feeling empathic concern under favorable conditions after having done so some number of times before, and so we may stop judging that we ought to save the next child that crosses our path after some number of children have already crossed our path.

Let me concede off the bat that many of us would likely stop feeling empathic concern after encountering some relatively small number of children, even if we encountered each child under favorable conditions. But as I briefly argued in chapter 2, collecting the evidence provided by empathic concern does not merely require us to feel empathic concern under favorable conditions. It also requires us to avoid burnout. So in fact premise 4 should read as follows:

(4*) We have decisive reason to believe that if we collected the evidence provided by empathic concern in the proper way and all else were equal, we would ask whether we should save one child (and then a second child, and then a third child...) under favorable conditions *without experiencing burnout*, and then infer

> that the cumulative number of children we end up judging we ought to save across these cases is the number of children we ought to save right now.

Why should we think that collecting the evidence provided by burnout requires us to avoid burnout? Because empathic concern provides us with evidence of our *pro tanto* reasons to help others. And I take it that all parties to the debate—both those who believe we should impoverish ourselves in order to help those in extreme poverty and those who believe we are required to give only a little to help those in extreme poverty—agree that our pro tanto reason to help the first drowning child we come across is just as strong as our pro tanto reason to help each child we come across thereafter. In each case, the reason to help is the bare fact that if we do not, an innocent child will die prematurely and be deprived of everything good their future would have to offer them. So if, as everyone agrees, our pro tanto reason to help the first drowning child is quite strong, and, as I have argued, the intense empathic concern we feel for that child gives us evidence of that, then we should feel equally intense empathic concern for every child thereafter if we want our judgment about whether we should help each child to be informed by all the high quality evidence bearing on that question. So it matters not at all whether we would actually reach burnout if we carried out the procedure of collecting empathic concern in the way I envision. What matters is whether idealized empathizers, versions of ourselves who were not subject to the mere physiological limitation of burnout, were to carry out this procedure.

One might wonder why, if all parties to the debate already agree that our pro tanto reasons to help each of the children are equally strong, we would need to collect the evidence provided by empathic concern. Why must we collect evidence for the truth of some proposition we already believe? I submit that we must still collect the evidence provided by empathic

concern. Why? Because only by collecting the evidence provided by empathic concern can our moral judgment take full account of the strength of our pro tanto reasons to help. Explicit beliefs about reasons can only influence our all-things-considered judgments in rare cases in which (e.g.) we already believe that reasons to perform an act and reasons against performing that act are perfectly balanced, and where it's obvious that the existence of additional reason on one side or the other could tip the balance. This may be in part because our explicit beliefs about reasons leave the strength of those reasons vague in a way that empathic concern does not. It may also be due in part to facts about the relative impenetrability of the processes that produce our intuitions. In any case, we know that experiences of empathic concern do influence people's judgments about whether they ought to help even in cases where they already took themselves to have strong reason to help. The best explanation of why this is that the evidence provided empathic concern can inform moral judgment in a way that explicit beliefs about reasons cannot.

At this stage, the objector may object to the original assumption that empathic concern gives us evidence of our *pro tanto* reasons. They may ask: why must we assume that empathic concern is giving us evidence of our *pro tanto* reasons to help and not harm rather than, say, the *balance* of our reasons to help and not harm? If empathic concern were merely giving us evidence regarding the balance of our reasons to help and not harm, then I would be begging the question in assuming those who avoid "burnout" enjoy access to non-misleading evidence the rest of us have reason to seek out. For I could only justify the claim that this is non-misleading evidence if I already believed that the balance of the reasons to help in each case fell in favor of helping. And that is exactly what I am intending to show.

To avoid this worry, I must be able to answer the question of why we should assume empathic concern only gives us evidence of pro tanto reasons. The answer is twofold.

First, as argued in the previous chapter, we have a firm intuition that empathic concern is a *fitting* response to the misfortune of others, and it does not seem any less fitting to feel empathic concern for the nth suffering individual than it does for the first individual, when both individuals have suffered equal misfortunes and are equally badly off. It would be strange if having the fitting response to someone's misfortune systematically gave us misleading evidence about our reasons to help or not harm them. Perhaps this happens in cases of disgust or envy—where feeling disgust or envy is fitting but does not provide us with any good evidence that we have reason to do the things disgust or envy dispose us to do—but this does not seem to be the case with empathic concern.

Second, if empathic concern were giving us evidence only of the *balance* of our reasons to help and not harm—of the strength of our obligation to help and not harm, all-things-considered—then we would expect that neurotypical individuals would be led astray by the empathic concern they feel during their moral development, for much of the empathic concern we learn from over the course of our moral development is felt for individuals who we should not help, all-things-considered. For example, we often feel empathic concern for others who have suffered some misfortune, but who we do not have any all-things-considered obligation to help (e.g., because doing so would be too costly for us). If feeling empathic concern gave us high quality evidence that we have an all-things-considered obligation to help the person we feel empathic concern for, then we would expect that we would treat it as though it provides this evidence. This would in turn lead us to believe we have all-things-considered obligations to help individuals whenever we feel empathic concern for them. But experiencing normal levels of empathic concern does not lead us to believe we have obligations that we do not have—far from it. If anything, feeling normal levels of empathic concern still leaves many of us underestimating

the strength of our all-things-considered obligations to help and not harm others. So we should assume that empathic concern is only giving us evidence of our pro tanto reasons to help and not harm.

Suppose this is all right. Empathic concern provides evidence regarding the strength of our *pro tanto* reasons to help, and collecting it in the proper way requires us to avoid burnout. So we should find some way to avoid burnout or defer to idealized versions of ourselves who were capable of avoiding burnout. The objector may grant all this and still reasonably wonder: how do we know what idealized versions of ourselves who escaped burnout would judge? While I believe what I initially said about the power of imagination to show us what we would judge holds, I will spend the remainder of this section offering additional empirical evidence to think that we would be much more charitable if we put ourselves in favorable conditions and avoided burnout.

No studies have been done on people who have followed the procedure I am recommending. Nevertheless, a body of work has recently begun to develop around individuals who are unusually selfless, individuals who have come to be called "extraordinary altruists". If it turned out that extraordinary altruists were extraordinarily altruist in part because they were especially good at getting themselves into favorable conditions for feeling empathic concern or avoided burnout in a way that most of us do not, this lends support for the view that we would judge we ought to save many children if we went through the procedure I recommend of feeling empathic concern for a series of children without experiencing burnout.

What do we know about extraordinary altruists? Most studies done thus far on extraordinary altruists focus on altruistic kidney donors: those who give kidneys to strangers who need them in return for no material goods. The evidence available suggests that these

extraordinary altruists have reduced psychopathic traits across the board, are unusually good at recognizing fear in other people's faces (which recall, is a favorable condition that psychopaths have trouble occupying), and have a correspondingly increased amygdala responses to other people's fear, which suggests an increase in empathic concern (Marsh et al. 2014). Indeed, independent evidence has found that this increased amygdala response in extraordinary altruists is associated with increased empathic concern (Brethel-Haurwitz et al. 2017). Finally, extraordinary altruists are also much less sensitive to the closeness bias described in chapter 2, the tendency to feel less empathic concern for those to whom has a less close relationship, and this lack of sensitivity is correlated with having traits correlated with higher empathic concern (Vekaria et al. 2017). Importantly, this is not because they extraordinary altruists believe themselves to be closer to others. This failure to be less generous to strangers suggests they enjoy a sort of immunity to a *kind* of burnout.

In addition to findings on altruistic kidney donors, a recent study done in a virtual reality environment found that those who were willing to risk their own lives to save someone in an emergency fire scenario were higher in empathic concern (Patil et al. 2017). This again suggests that those who are less susceptible to burnout take their obligations to help to be more demanding than others.

So far I have discussed extraordinarily altruistic humans. But we can also look to other species. And as it turns out, scientists have been able to induce extraordinary altruism in nonhuman animals, and these findings may be able to tell us something about what would happen to humans if they could avoid burnout. For example, if given the opportunity, pregnant rats will devote most of their time and energy to retrieving and caring for hundreds of unfamiliar rats (Wilsoncroft 1968). As Abigail Marsh puts it, one of the rat mothers retrieved 684 pups

during the experiment, "a pup every fifteen seconds for three hours, without a break, carted back to the nest for a total distance of over 2,000 feet," but that even the least altruistic mother retrieved 247 pups. (Marsh 2017; pp. 170-172) And while virgin female rats quickly go from completely ignoring unrelated rat pups to full maternal caretaking of multiple rats pups when injected with oxytocin (Pedersen & Prange 1989; Pedersen et al. 1982), a hormone shown to increase empathic concern and generosity in humans as well (Rodrigues et al. 2009; Zak et al. 2007; Hurlemann et al. 2010; Hubble et al. 2017; Barraza & Zak 2009). This provides some evidence that if we could increase empathic concern and prevent burnout in humans by injecting them with oxytocin, they would engage in repeated altruistic sacrifice as well.

I have argued that extraordinary altruists take their reasons to be stronger to help others to be stronger than the rest of us do because they avoid burnout, and we have reason to avoid burnout. But one may also hold that extraordinary altruists take their sacrifices to be morally required not only because they avoid burnout following repeated exposure to those in need, but because they feel more empathic concern even during their first encounter with someone in need. Not only does their empathic concern not diminish as quickly over iterated cases as the empathic concern of others, but they start off with a larger stock of empathic concern even in the first case. If this were true, we would need some reason for thinking that the greater empathic concern that extraordinary altruists initially experience even in a one-off case is providing them with higher quality evidence than the rest of us have access to. But why would this be?

I submit that if extraordinary altruists do feel greater empathic concern than the rest of us when they first encounter someone in need—before questions about burnout come into the picture—we should think that this greater empathic concern provides higher quality evidence than the evidence provided by the lesser empathic concern the rest of us experience. Why?

Because this greater empathic concern is felt under conditions that are more epistemically favorable than the lesser empathic concern. As suggested above, extraordinary altruists are especially good at getting themselves into epistemically favorable conditions—of detecting fear in other people's faces, and of falling victim to the closeness bias, of feeling more empathic concern for those who happen to be close to them in morally irrelevant ways. So if the reason extraordinary altruists take morality to be more demanding than the rest of us is that they feel more empathic concern even in one-off cases, we should still defer to them, because they seem to have access to high quality evidence the rest of us lack, they seem to share all of our other evidence, and they do not seem to be especially subject to any other epistemic defect.

The extraordinary altruists discussed so far—both human and nonhuman—are all willing to make exceptional sacrifices in order to bestow benefits that are often much smaller than the kinds of benefits we could provide by making much smaller sums to charity. Of course, the fact that such individuals are willing to make these sacrifices does not show that they believe they are *required* to. (This seems most unlikely especially in the case of rats.) I have argued that those who would avoid burnout in carrying out the serial empathy procedure I have recommended in cases like *Drowning Children* would take themselves to be required to save many children at a cumulatively large sacrifice to themselves. I have just provided evidence that those least likely to endure burnout do make significant sacrifices to help others. But I have not shown that they judge they *must*. And this is what would be required to vindicate premise (5).

Unfortunately, the studies performed so far have not asked extraordinary altruists whether they believe they were morally required to make the sacrifices they made. For all these studies shows, these altruists may believe their altruistic acts to be merely supererogatory. To gain evidence that those who would avoid burnout believe themselves morally required to make great

sacrifices, we must turn to anecdotes. And once we do, it becomes clear that the overall pattern is to think that extraordinary altruists believe their sacrifices to be morally required. Consider the familiar case of someone interviewed on the news who endangered their own life to save someone else from certain death. These individuals regularly tell the news reporter that they were just doing what any decent person would do, that they are not heroes, and so on and so forth. This strongly suggests they do not take their actions to be supererogatory. Moreover, those who do explicitly judge their own actions tend to think they are obligatory. Take, for example, Zell Kravinsky, who has publicly said that those who do not give away their kidney because of the chance of death to themselves place a value on their own life that is so much greater than the value they place on the lives of others so as to be considered "obscene" (Singer 2006). And consider what Larissa MacFarquhar writes in her close study of extraordinary altruists in *Strangers Drowning*:

> What do-gooders and utilitarians do have in common is the belief that most ordinary ideas about how much you owe to strangers, and how much duty you can demand of you, are wrong. (p. 121)

> She agonized over whether to go birthday parties, because she felt she couldn't show up without bringing a present, but she believed it would be wrong to spend five dollars on a present when that five dollars could be given to someone who needed it more. (p. 143)

> He kept thinking how wrong it was for them to have so much when so many others had nothing… (p. 57)

Extraordinary altruists are not merely doing what they take to be supererogatory. They take their sacrifices to be moral obligations. And they do not think they are special in having these obligations. We all have these obligations.

**Section 5: Conclusion**

I have argued that if we collected the evidence provided by empathic concern in the proper way, we would think that morality requires us to sacrifice much more to help those in need than we currently think. I have not said exactly how much it requires us to sacrifice—at the end of the day, figuring out exactly what we would judge if we collected empathic concern in the way I recommend would require us to actually collect empathic concern in this way, which would be quite difficult—but I have provided strong reason to think we are at least required to make sacrifices as much as real-life extraordinary altruists such as altruistic kidney donors.

## References


Aharoni, E., Antonenko, O., & Kiehl, K. A. (2011). Disparities in the moral intuitions of criminal offenders: The role of psychopathy. *Journal of Research in Personality*, *45*(3), 322–327. https://doi.org/10.1016/j.jrp.2011.02.005

Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. A. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *Journal of Abnormal Psychology*, *121*(2), 484–497. https://doi.org/10.1037/a0024796

Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. A. (2014). What's wrong? Moral understanding in psychopathic offenders. *Journal of Research in Personality*, *53*, 175–181. https://doi.org/10.1016/j.jrp.2014.10.002

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, *2*(11), 1032–1037. https://doi.org/10.1038/14833

Aniskiewicz, A. S. (1979). Autonomic components of vicarious conditioning and psychopathy. *Journal of Clinical Psychology*, *35*(1), 60–67.

Arbuckle, N. L., & Shane, M. S. (2017). Up-regulation of neural indicators of empathic concern in an offender population. *Social Neuroscience*, *12*(4), 386–390. https://doi.org/10.1080/17470919.2016.1179669

Ashar, Y. K., Andrews-Hanna, J. R., Dimidjian, S., & Wager, T. D. (2017). Empathic care and distress: Predictive brain markers and dissociable brain systems. *Neuron*, *94*(6), 1263–1273.e4. https://doi.org/10.1016/j.neuron.2017.05.014

Barbey, A. K., Colom, R., & Grafman, J. (2014). Neural mechanisms of discourse comprehension: a human lesion study. *Brain*, *137*(1), 277–287. https://doi.org/10.1093/brain/awt312

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(5), 1195–1205. https://doi.org/10.1016/j.cortex.2012.05.022

Barrash, J., Tranel, D., & Anderson, S. W. (2000). Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. *Developmental Neuropsychology*, *18*(3), 355–381. https://doi.org/10.1207/S1532694205Barrash

Barraza, J. A., & Zak, P. J. (2009). Empathy toward strangers triggers oxytocin release and subsequent generosity. *Annals of the New York Academy of Sciences*, *1167*, 182–189. https://doi.org/10.1111/j.1749-6632.2009.04504.x

Barraza, J. A., & Zak, P. J. (2009). Empathy toward strangers triggers oxytocin release and subsequent generosity. *Annals of the New York Academy of Sciences*, *1167*, 182–189. https://doi.org/10.1111/j.1749-6632.2009.04504.x

Barraza, J. A., Alexander, V., Beavin, L. E., Terris, E. T., & Zak, P. J. (2015). The heart of the story: peripheral physiology during narrative exposure predicts charitable giving. *Biological Psychology*, *105*, 138–143. https://doi.org/10.1016/j.biopsycho.2015.01.008

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*(1), 154–161. https://doi.org/10.1016/j.cognition.2011.05.010

Batson, C. D. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, *40*(2), 290–302.

Batson, C. D. (2011). *Altruism in Humans*. New York: Oxford University Press.

Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*, *3*(1), 141–177. https://doi.org/10.1111/j.1751-2409.2009.01013.x

Batson, C. D., Lishner, D. A., Cook, J., & Sawyer, S. (2005). Similarity and nurturance: Two possible sources of empathy for strangers. *Basic and Applied Social Psychology*, *27*(1), 15–25. https://doi.org/10.1207/s15324834basp2701_2

Beardsworth, A. D., & Keil, E. T. (1991). Vegetarianism, veganism, and meat avoidance: Recent trends and findings. *British Food Journal*, *93*(4), 19–24. https://doi.org/10.1108/00070709110135231

Beussink, C. N., Hackney, A. A., & Vitacco, M. J. (2017). The effects of perspective taking on empathy-related responses for college students higher in callous traits. *Personality and Individual Differences*, *119*(Supplement C), 86–91. https://doi.org/10.1016/j.paid.2017.07.005

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition*, *57*(1), 1–29.

Blair, R. J. R. (1997). Moral reasoning and the child with psychopathic tendencies. Personality and Individual Differences, 22(5), 731–739. https://doi.org/10.1016/S0191-8869(96)00249-8

Blair, R. J. R. (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. Journal of Neurology, Neurosurgery & Psychiatry, 71(6), 727–731. https://doi.org/10.1136/jnnp.71.6.727

Blair, R. J. R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, *11*(9), 387–392. https://doi.org/10.1016/j.tics.2007.07.003

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews. Neuroscience*, *14*(11), 786–799. https://doi.org/10.1038/nrn3577

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews. Neuroscience*, *14*(11), 786–799. https://doi.org/10.1038/nrn3577

Blair, R. J. R. (2017). Emotion-based learning systems and the development of morality. *Cognition*, *167*, 38–45. https://doi.org/10.1016/j.cognition.2017.03.013

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1995). Is the psychopath "morally insane"? *Personality and Individual Differences*, *19*(5), 741–752. https://doi.org/10.1016/0191-8869(95)00087-M

Blair, R. J., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: a lack of responsiveness to distress cues? *Psychophysiology*, *34*(2), 192–198.

Blair, R. J., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain: A Journal of Neurology*, *122* (5), 883–893.

Bloom, P. (2016). *Against Empathy: The Case for Rational Compassion*. New York, NY: Ecco.

Bornemann, B., Kok, B. E., Böckler, A., & Singer, T. (2016). Helping from the heart: Voluntary upregulation of heart rate variability predicts altruistic behavior. *Biological Psychology*, *119*, 54–63. https://doi.org/10.1016/j.biopsycho.2016.07.004

Brethel-Haurwitz, K. M., O'Connell, K., Cardinale, E. M., Stoianova, M., Stoycos, S. A., Lozier, L. M., … Marsh, A. A. (2017). Amygdala–midbrain connectivity indicates a role for the mammalian parental care system in human altruism. *Proc. R. Soc. B*, *284*(1865), 20171731. https://doi.org/10.1098/rspb.2017.1731

Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: a general deficit of interoception. *Royal Society Open Science*, *3*(10), 150664. https://doi.org/10.1098/rsos.150664

Brewer, R., Marsh, A. A., Catmur, C., Cardinale, E. M., Stoycos, S., Cook, R., & Bird, G. (2015). The impact of autism spectrum disorder and alexithymia on judgments of moral acceptability. *Journal of Abnormal Psychology*, *124*(3), 589–595. https://doi.org/10.1037/abn0000076

Brito, S. A. D., Viding, E., Kumari, V., Blackwood, N., & Hodgins, S. (2013). Cool and hot executive function impairments in violent offenders with antisocial personality disorder with and without psychopathy. *PLOS One*, *8*(6), e65566. https://doi.org/10.1371/journal.pone.0065566

Budhani, S., Richell, R. A., & Blair, R. J. R. (2006). Impaired reversal but intact acquisition: Probabilistic response reversal deficits in adult individuals with psychopathy. *Journal of Abnormal Psychology*, *115*(3), 552–558. https://doi.org/10.1037/0021-843X.115.3.552

Budolfson, M. (in press). The inefficacy objection to consequentialism, and the problem with the [Singer/Norcross/Kagan] expected consequences response. *Philosophical Studies*.

Caes, L., Uzieblo, K., Crombez, G., De Ruddere, L., Vervoort, T., & Goubert, L. (2012). Negative emotional responses elicited by the anticipation of pain in others: psychophysiological evidence. *The Journal of Pain: Official Journal of the American Pain Society*, *13*(5), 467–476. https://doi.org/10.1016/j.jpain.2012.02.003

Cardinale, E. M., & Marsh, A. A. (2015). Impact of psychopathy on moral judgments about causing fear and physical harm. *PLOS One*, *10*(5), e0125708. https://doi.org/10.1371/journal.pone.0125708

CBS/AP. (2017). Florida man's drowning recorded by group of teens, police say. Retrieved November 19, 2017, from https://www.cbsnews.com/news/cocoa-florida-jamel-dunn-drowning-recorded-group-teens-police/

Cheng, Y., Lin, C.-P., Liu, H.-L., Hsu, Y.-Y., Lim, K.-E., Hung, D., & Decety, J. (2007). Expertise modulates the perception of pain in others. *Current Biology: CB*, *17*(19), 1708–1713. https://doi.org/10.1016/j.cub.2007.09.020

Chopik, W. J., O'Brien, E., & Konrath, S. H. (2017). Differences in empathic concern and perspective taking across 63 countries. *Journal of Cross-Cultural Psychology*, *48*(1), 23–38. https://doi.org/10.1177/0022022116673910

Church, R. M. (1959). Emotional reactions of rats to the pain of others. *Journal of Comparative and Physiological Psychology*, *52*(2), 132–134.

Cima, M., Tonnaer, F., & Hauser, M. D. (2010). Psychopaths know right from wrong but don't care. *Social Cognitive and Affective Neuroscience*, *5*(1), 59–67. https://doi.org/10.1093/scan/nsp051

Coke, J. S., Batson, C. D., & Mcdavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology*, *36*(7), 752–766.

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences*, *107*(40), 17433–17438. https://doi.org/10.1073/pnas.1009396107

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, *7*(9), 415–423. https://doi.org/10.1016/S1364-6613(03)00197-9

Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: the aversion to harmful action. *Emotion (Washington, D.C.)*, *12*(1), 2–7. https://doi.org/10.1037/a0025071

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research*, *41*(2), 81–94.

Davis, M. H. (1994). *Empathy: A social psychological approach*. Chicago: Brown & Benchmark Publishers.

Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social Neuroscience*, *11*(1), 1–14. https://doi.org/10.1080/17470919.2015.1029593

Decety, J., Skelly, L., Keith J. Yoder, & Kiehl, K. A. (2014). Neural processing of dynamic emotional facial expressions in psychopaths. *Social Neuroscience*, *9*(1), 36–49. https://doi.org/10.1080/17470919.2013.866905

Decety, J., Yang, C.-Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: an event-related brain potential study. *NeuroImage*, *50*(4), 1676–1682. https://doi.org/10.1016/j.neuroimage.2010.01.025

Djeriouat, H., & Trémolière, B. (2014). The dark triad of personality and utilitarian moral judgment: The mediating role of honesty/humility and harm/care. *Personality and Individual Differences*, *67*, 11–16. https://doi.org/10.1016/j.paid.2013.12.026

Dombrowski, D. A. (1997). *Babies and beasts: The argument from marginal cases*. Urbana: University of Illinois Press.

Domes, G., Hollerbach, P., Vohs, K., Mokros, A., & Habermeyer, E. (2013). Emotional empathy and psychopathy in offenders: an experimental study. *Journal of Personality Disorders*, *27*(1), 67–84. https://doi.org/10.1521/pedi.2013.27.1.67

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology*, *6*, 495–525. https://doi.org/10.1146/annurev.clinpsy.121208.131208

Eslinger, P. J., Flaherty-Craig, C. V., & Benton, A. L. (2004). Developmental outcomes after early prefrontal cortex damage. *Brain and Cognition*, *55*(1), 84–103. https://doi.org/10.1016/S0278-2626(03)00281-1

Fairchild, G., van Goozen, S. H. M., Stollery, S. J., Aitken, M. R. F., Savage, J., Moore, S. C., & Goodyer, I. M. (2009). Decision making and executive function in male adolescents with early-onset or adolescence-onset conduct disorder and control subjects. *Biological Psychiatry*, *66*(2), 162–168. https://doi.org/10.1016/j.biopsych.2009.02.024

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *Neuroimage*, *105*, 347–356. https://doi.org/10.1016/j.neuroimage.2014.10.043

Finger, E. C., Marsh, A. A., Blair, K. S., Reid, M. E., Sims, C., Ng, P., … Blair, R. J. R. (2011). Disrupted reinforcement signaling in orbital frontal cortex and caudate in youths with conduct disorder/oppositional defiant disorder and high psychopathic traits. *The American Journal of Psychiatry*, *168*(2), 152–162. https://doi.org/10.1176/appi.ajp.2010.10010129

Fisher, L., & Blair, R. J. (1998). Cognitive impairment and its relationship to psychopathic tendencies in children with emotional and behavioral difficulties. *Journal of Abnormal Child Psychology*, *26*(6), 511–519.

Foulkes, L., McCrory, E. J., Neumann, C. S., & Viding, E. (2014). Inverted social reward: Associations between psychopathic traits and self-report and experimental measures of social reward. *PLOS One*, *9*(8). https://doi.org/10.1371/journal.pone.0106000

Fox, N., & Ward, K. (2008). Health, ethics and environment: a qualitative study of vegetarian motivations. *Appetite*, *50*(2-3), 422–429. https://doi.org/10.1016/j.appet.2007.09.007

Francis, K. B., Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., & Terbeck, S. (2016). Virtual morality: Transitioning from moral judgment to moral action? *PLOS One*, *11*(10), e0164374. https://doi.org/10.1371/journal.pone.0164374

Francis, L. P., & Norman, R. (1978). Some animals are more equal than others. *Philosophy*, *53*(206), 507–527. https://doi.org/10.1017/S0031819100026358

Frey, R. G. (1980). *Interests and rights: the case against animals*. Oxford University Press, Incorporated.

Frick, P. J. (1998). Callous-unemotional traits and conduct problems: Applying the two-factor model of psychopathy to children. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), *Psychopathy: theory, research and implications for society* (pp. 161–187). Dordrecht: Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-94-011-3965-6_8

Fultz, J., Batson, C. D., Fortenbach, V. A., McCarthy, P. M., & Varney, L. L. (1986). Social evaluation and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology*, *50*(4), 761–769.

Gao, Y., & Tang, S. (2013). Psychopathic personality and utilitarian moral judgment in college students. *Journal of Criminal Justice*, *41*, 342–349.

Gert, B. (2005). *Morality: Its nature and justification* (Rev. ed.). New York, NY: Oxford University Press.

Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PLOS One*, *8*(4), e60418. https://doi.org/10.1371/journal.pone.0060418

Glenn, A. L., Iyer, R., Graham, J., Koleva, S., & Haidt, J. (2009a). Are all types of morality compromised in psychopathy? *Journal of Personality Disorders*, *23*(4), 384–398. https://doi.org/10.1521/pedi.2009.23.4.384

Glenn, A. L., Raine, A., & Schug, R. A. (2009). The neural correlates of moral decision-making in psychopathy. *Molecular Psychiatry*, *14*(1), 5–6. https://doi.org/10.1038/mp.2008.104

Glenn, A. L., Raine, A., Schug, R. A., Young, L., & Hauser, M. (2009b). Increased DLPFC activity during moral decision-making in psychopathy. *Molecular Psychiatry*, *14*(10), 909–911. https://doi.org/10.1038/mp.2009.76

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Gur, R. C., & Sachser, N. (2009). Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults. *Ethology*, *115*(3), 257–263. https://doi.org/10.1111/j.1439-0310.2008.01603.x

Goel, V., & Dolan, R. J. (2003). Reciprocal neural response within lateral and ventral medial prefrontal cortex during hot and cold reasoning. *NeuroImage*, *20*(4), 2314–2321.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Greene, J. D. (2017). The rat-a-gorical imperative: Moral intuition and the limits of affective learning. *Cognition*, *167*, 66–77. https://doi.org/10.1016/j.cognition.2017.03.004

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Greene, J. T. (1969). Altruistic behavior in the albino rat. *Psychonomic Science*, *14*(1), 47–48. https://doi.org/10.3758/BF03336420

Hare, R. D. (1999). *Without conscience: The disturbing world of the psychopaths among us*. New York, NY: Guilford Press.

Harman. E. (2015). Eating meat as a morally permissible moral mistake. In A. Chignell., T. Cuneo, and M. Haltemann (Eds.), *Philosophy comes to dinner* (pp. 215-231). New York, NY: Routledge.

Haslam, N., Loughnan, S., & Perry, G. (2014). Meta-Milgram: An empirical synthesis of the obedience experiments. *PLOS One*, *9*(4), e93927. https://doi.org/10.1371/journal.pone.0093927

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149–160. https://doi.org/10.1016/j.neuron.2010.09.003

Hein, G., Engelmann, J. B., Vollberg, M. C., & Tobler, P. N. (2016). How learning shapes the empathic brain. *Proceedings of the National Academy of Sciences*, *113*(1), 80–85. https://doi.org/10.1073/pnas.1514539112

Hiraoka, D., & Nomura, M. (2016). The influence of cognitive load on empathy and intention in response to infant crying. *Scientific Reports*, *6*, srep28247. https://doi.org/10.1038/srep28247

Hiraoka, D., & Nomura, M. (2017). Would situational stress be harmful for anyone? The influence of situational factors and trait empathy on women's response to infant crying. *Infant Behavior & Development*, *48*(Pt B), 147–156. https://doi.org/10.1016/j.infbeh.2017.04.005

House, T. H., & Milligan, W. L. (1976). Autonomic responses to modeled distress in prison psychopaths. *Journal of Personality and Social Psychology*, *34*(4), 556–560.

Hubble, K., Daughters, K., Manstead, A. S. R., Rees, A., Thapar, A., & van Goozen, S. H. M. (2017). Oxytocin increases attention to the eyes and selectively enhances self-reported affective empathy for fear. *Neuropsychologia*, *106*, 350–357. https://doi.org/10.1016/j.neuropsychologia.2017.10.019

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., … Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience*, *30*(14), 4999–5007. https://doi.org/10.1523/jneurosci.5538-09.2010

Hussar, K. M., & Harris, P. L. (2010). Children who choose not to eat meat: A study of early moral decision-making. *Social Development*, *19*(3), 627–641. https://doi.org/10.1111/j.1467-9507.2009.00547.x

James, S. M. (2007). Good Samaritans, good humanitarians. *Journal of Applied Philosophy*, *24*(3), 238–254. https://doi.org/10.1111/j.1468-5930.2007.00378.x

Jenni, K., & Loewenstein, G. (1997). Explaining the Identifiable Victim Effect. *Journal of Risk and Uncertainty*, *14*(3), 235–257. https://doi.org/10.1023/A:1007740225484

Johnston, B. M., & Glasford, D. E.. (2017). Intergroup contact and helping: How quality contact and empathy shape outgroup helping. *Group Processes & Intergroup Relations*, 1368430217711770. https://doi.org/10.1177/1368430217711770

Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior*, *9*(5), 441–476.

Kagan, S. (2011). Do I make a difference? *Philosophy & Public Affairs*, *39*(2), 105–141. https://doi.org/10.1111/j.1088-4963.2011.01203.x

Kagan, S. (2016). What's wrong with speciesism? *Journal of Applied Philosophy*, *33*(1), 1–21.

Kahane, G., Everett, J. A. C., Earp, B. D., Farias, M., & Savulescu, J. (2015). "Utilitarian" judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193–209. https://doi.org/10.1016/j.cognition.2014.10.005

Kamm, F. (1999). Famine ethics: The problem of distance in morality and Singer's ethical theory. In D. Jamieson (Ed.), *Singer and his critics* (pp. 174–203). Oxford: Blackwell.

Kamm, F. M. (2000). Does distance matter morally to the duty to rescue? *Law and Philosophy*, *19*(6), 655–681. https://doi.org/10.2307/3505070

Kennett, J. and Fine, C. (2008). Internalism and the evidence from psychopaths and 'acquired sociopaths'". In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 3: The neuroscience of morality* (pp. 173-190). Cambridge, MA: MIT Press.

Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopath: History, neuroscience, treatment, and economics. *Jurimetrics*, *51*, 355–397.

Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, *7*(6), 708–714. https://doi.org/10.1093/scan/nsr048

Kogut, T., & Ritov, I. (2005a). The "identified victim" effect: an identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18*(3), 157–167. https://doi.org/10.1002/bdm.492

Kogut, T., & Ritov, I. (2005b). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, *97*(2), 106–116. https://doi.org/10.1016/j.obhdp.2005.02.003

Kringelbach, M. L., Stark, E. A., Alexander, C., Bornstein, M. H., & Stein, A. (2016). On cuteness: unlocking the parental brain and beyond. *Trends in Cognitive Sciences*, *20*(7), 545–558. https://doi.org/10.1016/j.tics.2016.05.003

Kuehne, M., Heimrath, K., Heinze, H.-J., & Zaehle, T. (2015). Transcranial direct current stimulation of the left dorsolateral prefrontal cortex shifts preference of moral judgments. *PLOS One*, *10*(5), e0127061. https://doi.org/10.1371/journal.pone.0127061

Kumar, V. (2016). Psychopathy and internalism. *Canadian Journal of Philosophy*, *46*(3), 318–345. https://doi.org/10.1080/00455091.2016.1165571

Kunst, J. R., & Hohle, S. M. (2016). Meat eaters by dissociation: How we present, prepare and talk about meat increases willingness to eat meat by reducing empathy and disgust. *Appetite*, *105*, 758–774. https://doi.org/10.1016/j.appet.2016.07.009

Kunst, J. R., & Palacios Haugestad, C. A. (2018). The effects of dissociation on willingness to eat meat are moderated by exposure to unprocessed meat: A cross-cultural demonstration. *Appetite*, *120*(Supplement C), 356–366. https://doi.org/10.1016/j.appet.2017.09.016

Kurzban, R., DeScioli, P., & Fein, D. (2012). Hamilton vs. Kant: pitting adaptations for altruism against adaptations for moral judgment. *Evolution and Human Behavior*, *33*(4), 323–333. https://doi.org/10.1016/j.evolhumbehav.2011.11.002

Lang, S., Yu, T., Markl, A., Müller, F., & Kotchoubey, B. (2011). Hearing others' pain: neural activity related to empathy. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(3), 386. https://doi.org/10.3758/s13415-011-0035-0

Langlois, J. H., Ritter, J. M., Casey, R. J., & Sawin, D. B. (1995). Infant attractiveness predicts maternal behaviors and attitudes. *Developmental Psychology*, *31*(3), 464.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

Light, S. N., Moran, Z. D., Swander, L., Le, V., Cage, B., Burghy, C., … Davidson, R. J. (2015). Electromyographically assessed empathic concern and empathic happiness predict increased prosocial behavior in adults. *Biological Psychology*, *104*, 116–129. https://doi.org/10.1016/j.biopsycho.2014.11.015

Lishner, D. A., Hong, P. Y., Jiang, L., Vitacco, M. J., & Neumann, C. S. (2015). Psychopathy, narcissism, and borderline personality: A critical test of the affective empathy-impairment hypothesis. *Personality and Individual Differences*, *86*, 257–265. https://doi.org/10.1016/j.paid.2015.05.036

Lishner, D. A., Oceja, L. V., Stocks, E. L., & Zaspel, K. (2008). The effect of infant-like characteristics on empathic concern for adults in need. *Motivation and Emotion*, *32*(4), 270–277. https://doi.org/10.1007/s11031-008-9101-5

Lishner, D. A., Vitacco, M. J., Hong, P. Y., Mosley, J., Miska, K., & Stocks, E. L. (2012). Evaluating the relation between psychopathy and affective empathy: two preliminary studies. *International Journal of Offender Therapy and Comparative Criminology*, *56*(8), 1161–1181. https://doi.org/10.1177/0306624X11421891

Lockwood, P. L., Hamonet, M., Zhang, S. H., Ratnavel, A., Salmony, F. U., Husain, M., & Apps, M. A. J. (2017). Prosocial apathy for helping others when effort is required. *Nature Human Behaviour*, *1*(7), s41562–017–0131–017. https://doi.org/10.1038/s41562-017-0131

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

Lovett, B. J., & Sheffield, R. A. (2007). Affective empathy deficits in aggressive children and adolescents: a critical review. *Clinical Psychology Review*, *27*(1), 1–13. https://doi.org/10.1016/j.cpr.2006.03.003

Lui, J. H. L., Barry, C. T., & Sacco, D. F. (2016). Callous-unemotional traits and empathy deficits: Mediating effects of affective perspective-taking and facial emotion recognition. *Cognition & Emotion*, *30*(6), 1049–1062. https://doi.org/10.1080/02699931.2015.1047327

MacFarquhar, L. (2016). *Strangers drowning: Impossible idealism, drastic choices, and the urge to help* (Reprint edition). New York, New York: Penguin Books.

Marsh, A. (2017). *The fear factor: How one emotion connects altruists, psychopaths, and everyone in-between*. New York, NY: Basic Books.

Marsh, A. A., & Cardinale, E. M. (2012). Psychopathy and fear: specific impairments in judging behaviors that frighten others. *Emotion (Washington, D.C.)*, *12*(5), 892–898. https://doi.org/10.1037/a0026260

Marsh, A. A., & Cardinale, E. M. (2014). When psychopathy impairs moral judgments: neural responses during judgments about causing fear. *Social Cognitive and Affective Neuroscience*, *9*(1), 3–11. https://doi.org/10.1093/scan/nss097

Marsh, A. A., Stoycos, S. A., Brethel-Haurwitz, K. M., Robinson, P., VanMeter, J. W., & Cardinale, E. M. (2014). Neural and cognitive characteristics of extraordinary altruists. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), 15036–15041. https://doi.org/10.1073/pnas.1408440111

Marshall, J., Watts, A. L., & Lilienfeld, S. O. (2016). Do psychopathic individuals possess a misaligned moral compass? A meta-analytic examination of psychopathy's relations with moral judgment. *Personality Disorders*. https://doi.org/10.1037/per0000226

Marshall, J., Watts, A. L., Frankel, E. L., & Lilienfeld, S. O. (2017). An examination of psychopathy's relationship with two indices of moral judgment. *Personality and Individual Differences*, *113*, 240–245. https://doi.org/10.1016/j.paid.2017.03.034

Masserman, J. H., Wechkin, S., & Terris, W. (1964). "Altruistic" behavior in rhesus monkeys. *The American Journal of Psychiatry*, *121*, 584–585. https://doi.org/10.1176/ajp.121.6.584

McDonald, M. M., Defever, A. M., & Navarrete, C. D. (2017). Killing for the greater good: Action aversion and the emotional inhibition of harm in moral dilemmas. *Evolution and Human Behavior*, *0*(0). https://doi.org/10.1016/j.evolhumbehav.2017.06.001

Meffert, H., Gazzola, V., den Boer, J. A., Bartels, A. A. J., & Keysers, C. (2013). Reduced spontaneous but relatively normal deliberate vicarious representations in psychopathy. *Brain: A Journal of Neurology*, *136*(Pt 8), 2550–2562. https://doi.org/10.1093/brain/awt190

Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, *103*(3), 324–344.

Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion (Washington, D.C.)*, *14*(3), 573–587. https://doi.org/10.1037/a0035361

Miller, R. W. (2004). Beneficence, duty and distance. *Philosophy and Public Affairs*, *32*(4), 357–383.

Miller, R., & Cushman, F. (2013). Aversive for me, wrong for you: First-person behavioral aversions underlie the moral condemnation of harm. *Social and Personality Psychology Compass*, *7*(10), 707–718. https://doi.org/10.1111/spc3.12066

Morelli, S. A., Rameson, L. T., & Lieberman, M. D. (2014). The neural components of empathy: predicting daily prosocial behavior. *Social Cognitive and Affective Neuroscience*, *9*(1), 39–47. https://doi.org/10.1093/scan/nss088

Narveson, J. (2003). We don't owe them a thing! A tough-minded but soft-hearted view of aid to the faraway needy. *The Monist*, *86*(3), 419–433.

Nefsky, J. (2011). Consequentialism and the problem of collective harm: A reply to Kagan. *Philosophy & Public Affairs*, *39*(4), 364–395. https://doi.org/10.1111/j.1088-4963.2012.01209.x

Nefsky, J. (2017). How you can help, without making a difference. *Philosophical Studies*, *174*(11), 2743–2767. https://doi.org/10.1007/s11098-016-0808-y

Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*, *95*(3), 252–256.

Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. New York, NY: Oxford University Press.

Norcross, A. (2004). Puppies, pigs, and people: Eating meat and marginal cases. *Philosophical Perspectives*, *18*(1), 229–245. https://doi.org/10.1111/j.1520-8583.2004.00027.x

Nucci, L. P. (2001). *Education in the moral domain*. Cambridge, UK: Cambridge University Press.

O'Brien, B. S., & Frick, P. J. (1996). Reward dominance: associations with anxiety, conduct problems, and psychopathy in children. *Journal of Abnormal Child Psychology*, *24*(2), 223–240.

Oliver, L. D., Neufeld, R. W. J., Dziobek, I., & Mitchell, D. G. V. (2016). Distinguishing the relationship between different aspects of empathic responding as a function of psychopathic, autistic, and anxious traits. *Personality and Individual Differences*, *99*, 81–88. https://doi.org/10.1016/j.paid.2016.04.058

Parfit. D. (2017). *On what matters: Volume three*. New York, NY: Oxford University Press.

Patil, I. (2015). Trait psychopathy and utilitarian moral judgement: The mediating role of action aversion. *Journal of Cognitive Psychology*, *27*(3), 349–366. https://doi.org/10.1080/20445911.2015.1004334

Patil, I., & Silani, G. (2014a). Alexithymia increases moral acceptability of accidental harms. *Journal of Cognitive Psychology*, *26*(5), 597–614. https://doi.org/10.1080/20445911.2014.929137

Patil, I., & Silani, G. (2014b). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00501

Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social Neuroscience*, *9*(1), 94–107. https://doi.org/10.1080/17470919.2013.870091

Patil, I., Melsbach, J., Hennig-Fast, K., & Silani, G. (2016). Divergent roles of autistic and alexithymic traits in utilitarian moral judgments in adults with autism. *Scientific Reports*, *6*, srep23637. https://doi.org/10.1038/srep23637

Patil, I., Zanon, M., Novembre, G., Zangrando, N., Chittaro, L., & Silani, G. (2017). Neuroanatomical basis of concern-based altruism in virtual environment. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2017.02.015

Pedersen, C. A., & Prange, A. J. (1979). Induction of maternal behavior in virgin rats after intracerebroventricular administration of oxytocin. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(12), 6661–6665.

Pedersen, C. A., Ascher, J. A., Monroe, Y. L., & Prange, A. J. (1982). Oxytocin induces maternal behavior in virgin female rats. *Science*, *216*(4546), 648–650.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922–934. https://doi.org/10.1002/ejsp.504

Pfabigan, D. M., Seidel, E.-M., Wucherer, A. M., Keckeis, K., Derntl, B., & Lamm, C. (2015). Affective empathy differs in male violent offenders with high- and low-trait psychopathy. *Journal of Personality Disorders*, *29*(1), 42–61. https://doi.org/10.1521/pedi_2014_28_145

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron*, *43*(6), 897–905. https://doi.org/10.1016/j.neuron.2004.08.042

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., … Gray, J. A. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings. Biological Sciences*, *265*(1408), 1809–1817. https://doi.org/10.1098/rspb.1998.0506

Pletti, C., Lotto, L., Tasso, A., & Sarlo, M. (2016). Will I regret it? Anticipated negative emotions modulate choices in moral dilemmas. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01918

Pogge, T. W. (2008). *World poverty and human rights*. Polity.

Pollock, J. L. (1986). *Contemporary theories of knowledge*. London: Hutchinson.

Pribis, P., Pencak, R. C., & Grajales, T. (2010). Beliefs and attitudes toward vegetarian lifestyle across generations. *Nutrients*, *2*(5), 523–531. https://doi.org/10.3390/nu2050523

Prinz, J. (2007). *The emotional construction of morals*. Oxford: Oxford University Press.

Prinz, J. (2011). Against empathy. *The Southern Journal of Philosophy*, *49*, 214–233. https://doi.org/10.1111/j.2041-6962.2011.00069.x

Quas, J. A., Dickerson, K. L., Matthew, R., Harron, C., & Quas, C. M. (2017). Adversity, emotion recognition, and empathic concern in high-risk youth. *PLOS One*, *12*(7), e0181606. https://doi.org/10.1371/journal.pone.0181606

Quinn, W. S. (1990). The puzzle of the self-torturer. *Philosophical Studies*, *59*(1), 79–90.

Ritchie, M. B., & Forth, A. E. (2016). Without concern: Predicting personal-moral transgressions from psychopathy and gender. *Personality and Individual Differences*, *94*, 247–252. https://doi.org/10.1016/j.paid.2016.01.041

Robinson, E. V., & Rogers, R. (2015). Empathy faking in psychopathic offenders: The vulnerability of empathy measures. *Journal of Psychopathology and Behavioral Assessment*, *37*(4), 545–552. https://doi.org/10.1007/s10862-015-9479-9

Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences*, *106*(50), 21437–21441. https://doi.org/10.1073/pnas.0909579106

Rothgerber, H. (2013). Real men don't eat (vegetable) quiche: Masculinity and the justification of meat consumption. *Psychology of Men & Masculinity*, *14*, 363. https://doi.org/10.1037/a0030379

Rothgerber, H. (2015a). Can you have your meat and eat it too? Conscientious omnivores, vegetarians, and adherence to diet. *Appetite*, *84*, 196–203. https://doi.org/10.1016/j.appet.2014.10.012

Rothgerber, H. (2015b). Underlying differences between conscientious omnivores and vegetarians in the evaluation of meat and animals. *Appetite*, *87*, 251–258. https://doi.org/10.1016/j.appet.2014.12.206

Rothgerber, H., & Mican, F. (2014). Childhood pet ownership, attachment to pets, and subsequent meat avoidance. The mediating role of empathy toward animals. *Appetite*, *79*, 11–17. https://doi.org/10.1016/j.appet.2014.03.032

Rozin, P., Markwith, M., & Stoess, C. (1997). Moralization and becoming a begetarian: The transformation of preferences into values and the recruitment of disgust. *Psychological Science*, *8*(2), 67–73. https://doi.org/10.1111/j.1467-9280.1997.tb00685.x

Sakai, J. T., Dalwani, M. S., Gelhorn, H. L., Mikulich-Gilbertson, S. K., & Crowley, T. J. (2012). A behavioral test of accepting benefits that cost others: associations with conduct problems and callous-unemotionality. *PLOS One*, *7*(4), e36158. https://doi.org/10.1371/journal.pone.0036158

Sakai, J. T., Dalwani, M. S., Mikulich-Gilbertson, S. K., McWilliams, S. K., Raymond, K. M., & Crowley, T. J. (2016). A behavioral measure of costly helping: replicating and extending the association with callous unemotional traits in male adolescents. *PLOS One*, *11*(3), e0151678. https://doi.org/10.1371/journal.pone.0151678

Sakai, J. T., Dalwani, M. S., Mikulich-Gilbertson, S. K., Raymond, K., McWilliams, S., Tanabe, J., Rojas, D., Regner, M., Banich, M. T., & Crowley, T. J. (2017). Imaging decision about whether to benefit self by harming others: Adolescents with conduct and substance problems, with or without callous-unemotionality, or developing typically. *Psychiatry Research*, *263*, 103–112. https://doi.org/10.1016/j.pscychresns.2017.03.004

Sarlo, M., Lotto, L., Rumiati, R., & Palomba, D. (2014). If it makes you feel bad, don't do it! Egoistic rather than altruistic empathy modulates neural and behavioral responses in moral dilemmas. *Physiology & Behavior*, *130*, 127–134. https://doi.org/10.1016/j.physbeh.2014.04.002

Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, *29*(12), 1241–1249.

Seara-Cardoso, A., & Viding, E. (2015). Functional neuroscience of psychopathic personality in adults. *Journal of Personality*, *83*(6), 723–737. https://doi.org/10.1111/jopy.12113

Seara-Cardoso, A., Dolberg, H., Neumann, C., Roiser, J. P., & Viding, E. (2013). Empathy, morality and psychopathic traits in women. *Personality and Individual Differences*, *55*(3), 328–333.

Seara-Cardoso, A., Neumann, C., Roiser, J., McCrory, E., & Viding, E. (2012). Investigating associations between empathy, morality and psychopathic personality traits in the general population. *Personality and Individual Differences*, *52*(1), 67–71. https://doi.org/10.1016/j.paid.2011.08.029

Seara-Cardoso, A., Sebastian, C. L., Viding, E., & Roiser, J. P. (2016). Affective resonance in response to others' emotional faces varies with affective ratings and psychopathic traits in amygdala and anterior insula. *Social Neuroscience*, *11*(2), 140–152.

Seara-Cardoso, A., Viding, E., Lickley, R. A., & Sebastian, C. L. (2015). Neural responses to others' pain vary with psychopathic traits in healthy adult males. *Cognitive, Affective & Behavioral Neuroscience*, *15*(3), 578–588. https://doi.org/10.3758/s13415-015-0346-7

Schmidtz, D. (2000). Islands in a sea of obligation: Limits of the duty to rescue. *Law and Philosophy*, *19*(6), 683–705. https://doi.org/10.2307/3505071

Singer, P. (1972). Famine, affluence, and morality. *Philosophy & Public Affairs*, *1*(3), 229-243.

Singer, P. (1974). Sidgwick and reflective equilibrium. *The Monist*, *58*(3), 490–517.

Singer, P. (2006, December 17). What should a billionaire give – and what should you? *The New York Times*. Retrieved from https://www.nytimes.com/2006/12/17/magazine/17charity.t.html

Singer, P. (2009). *The Life You Can Save: Acting Now to End World Poverty*. Random House Publishing Group.

Stocks, E. L. (2001). Self-other merging and empathic concern: Has the egoism-altruism debate been resolved? Unpublished master's thesis, University of Kansas, Lawrence.

Taber-Thomas, B. C., Asp, E. W., Koenigs, M., Sutterer, M., Anderson, S. W., & Tranel, D. (2014). Arrested development: early prefrontal lesions impair the maturation of moral judgement. *Brain*, *137*(4), 1254–1261. https://doi.org/10.1093/brain/awt377

Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00250

Timmerman, T. (2015). Sometimes there is nothing wrong with letting a child drown. *Analysis*, *75*(2), 204–212. https://doi.org/10.1093/analys/anv015

Toi, M., & Batson, C. D. (1982). More evidence that empathy is a source of altruistic motivation. *Journal of Personality and Social Psychology*, *43*(2), 281.

Tolstoy, L. (2010). *A calendar of wisdom: daily thoughts to nourish the soul*. New York, NY: Simon and Schuster.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.

Tusche, A., Böckler, A., Kanske, P., Trautwein, F.-M., & Singer, T. (2016). Decoding the charitable brain: Empathy, perspective taking, and attention shifts differentially predict altruistic giving. *Journal of Neuroscience*, *36*(17), 4719–4732. https://doi.org/10.1523/JNEUROSCI.3392-15.2016

Unger, P. K. (1996). *Living high and letting die: Our illusion of innocence*. New York, NY: Oxford University Press.

Vachon, D. D., Lynam, D. R., & Johnson, J. A. (2014). The (non)relation between empathy and aggression: surprising results from a meta-analysis. *Psychological Bulletin*, *140*(3), 751–773. https://doi.org/10.1037/a0035236

Van Langen, M. A. M., Wissink, I. B., van Vugt, E. S., Van der Stouwe, T., & Stams, G. J. J. M. (2014). The relation between empathy and offending: A meta-analysis. *Aggression and Violent Behavior*, *19*(2), 179–189. https://doi.org/10.1016/j.avb.2014.02.003

Västfjäll, D., Slovic, P., Mayorga, M., & Peters, E. (2014). Compassion fade: Affect and charity are greatest for a single child in need. *PLOS One*, *9*(6). https://doi.org/10.1371/journal.pone.0100115

Vekaria, K. M., Brethel-Haurwitz, K. M., Cardinale, E. M., Stoycos, S. A., & Marsh, A. A. (2017). Social discounting and distance perceptions in costly altruism. *Nature Human Behaviour*, *1*(5), s41562–017–0100–017. https://doi.org/10.1038/s41562-017-0100

Vyas, K., Jameel, L., Bellesi, G., Crawford, S., & Channon, S. (2017). Derailing the trolley: Everyday utilitarian judgments in groups high versus low in psychopathic traits or autistic traits. *Psychiatry Research*, *250*, 84–91. https://doi.org/10.1016/j.psychres.2017.01.054

Watanabe, S., & Ono, K. (1986). An experimental analysis of "empathic" response: Effects of pain reactions of pigeon upon other pigeon's operant behavior. *Behavioural Processes*, *13*(3), 269–277. https://doi.org/10.1016/0376-6357(86)90089-6

Wechkin, S., Masserman, J. H., & Terris, W. (1964). Shock to a conspecific as an aversive stimulus. *Psychonomic Science*, *1*(1-12), 47–48. https://doi.org/10.3758/BF03342783

White, B. A. (2014). Who cares when nobody is watching? Psychopathic traits and empathy in prosocial behaviors. *Personality and Individual Differences*, *56*(Supplement C), 116–121. https://doi.org/10.1016/j.paid.2013.08.033

Wiech, K., Kahane, G., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2013). Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. *Cognition*, *126*(3), 364–372. https://doi.org/10.1016/j.cognition.2012.11.002

Wilsoncroft, W. E. (1968). Babies by bar-press: Maternal behavior in the rat. *Behavior Research Methods & Instrumentation*, *1*(6), 229–230. https://doi.org/10.3758/BF03208105

Winter, K., Spengler, S., Bermpohl, F., Singer, T., & Kanske, P. (2017). Social cognition in aggressive offenders: Impaired empathy, but intact theory of mind. *Scientific Reports*, *7*(1), 670. https://doi.org/10.1038/s41598-017-00745-0

Young, L., Koenigs, M., Kruepke, M., & Newman, J. P. (2012). Psychopathy Increases perceived moral permissibility of accidents. *Journal of Abnormal Psychology*, *121*(3), 659–667. https://doi.org/10.1037/a0027489

Your Dollar Goes Further Overseas. (n.d.). Retrieved November 19, 2017, from https://www.givewell.org/giving101/Your-dollar-goes-further-overseas

Zak, P. J., Stanton, A. A., & Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLOS One*, 2(11), e1128. https://doi.org/10.1371/journal.pone.0001128

Zickfeld, J. H., Kunst, J. R., & Hohle, S. M. (2018). Too sweet to eat: Exploring the effects of cuteness on meat consumption. *Appetite*, *120*(Supplement C), 181–195. https://doi.org/10.1016/j.appet.2017.08.038

Zych, I., Ttofi, M. M., & Farrington, D. P. (2016). Empathy and callous–unemotional traits in different bullying roles. *Trauma, Violence, & Abuse*, 1524838016683456. https://doi.org/10.1177/1524838016683456