

The Moral Significance of Empathy



William Jefferson

Department of Philosophy
University of Oxford

This dissertation is submitted for the degree of
Doctor of Philosophy

Word count: 60,976

Balliol College

January 2019

Abstract

In this thesis, I argue that empathy is morally significant because it plays an important role in informing our moral deliberations. Empathy should be thought of not as an alternative to rational deliberation about how we are to act, but rather as an important input into such deliberation.

I focus on exploring what we learn when we empathize with the suffering of another person. Standard epistemic defences of empathy say only that such empathy will give us knowledge of which affective states the suffering person is feeling. I add to those defences by arguing that empathy with a suffering person also gives us additional types of knowledge about the affective states that this person is feeling. Most significantly, I argue that empathizing with a suffering person gives us knowledge of the strength of one of our reasons to help that person. I call this the Normative Epistemic Claim.

In chapter 1, I contextualize my approach within the recent philosophical discussion of empathy. In the following chapter, I explore the relationship between empathy and altruism, and argue that there is an unanswered question about how empathy gives rise to altruistic motivation. I suggest that we should answer this question by considering the idea that empathy gives us phenomenal knowledge about the affective states of other people. I defend this idea in chapter 3, where I also relate it to the debate about phenomenal knowledge that has been stimulated by Frank Jackson's example of Mary, the scientist who comes to see colour for the first time. In chapter 4, I begin to show the importance of phenomenal knowledge to moral deliberation. I argue that by giving us such knowledge, empathy with a suffering person also gives us knowledge of how

intrinsically bad that person's suffering is for them. In chapter 5, I extend this approach to defend the Normative Epistemic Claim. Finally, in chapter 6, I summarize the role of empathy in moral deliberation and broaden my discussion to include consideration of the significance that empathy has in inhibiting harming and in directing us to promote the joy of other people.

Table of Contents

Chapter 1: Introduction	7
1. Overview	7
2. Empathy	9
2.1 Affective Matching	9
2.2 Perspective-Taking	15
2.3 Affective Contagion	18
3. Against Empathy	24
3.1 Bucking the trend	24
3.2 Morality without empathy	25
3.3 The Emotional Cost Objection	27
3.4 The Spotlight Effect Objection	29
4. Defending Empathy	30
4.1 Selective Empathy	30
4.2 The Epistemic Defence	32
4.3 The Motivational Defence	34
4.4 The Relational Defence	35
5. Two Challenges for the Epistemic Defence	37
5.1 What is special about empathy?	37
5.2 The Spotlight Effect Objection Revisited	39
6. The Normative Epistemic Argument	45
7. The Plan for this Thesis	50
Chapter 2: Empathy and Altruism	52
1. Overview	52
2. The Explanatory Question	53
2.1 Steven and Oscar	53
2.2 Empathy as Affective Matching	54
2.3 The Empathy-Altruism Hypothesis	56
2.4 The Explanatory Question	59
3. Three Attempts to Answer the Explanatory Question	61
3.1 The Standard Epistemic Explanation	61

3.2 The Aversive Arousal Reduction Explanation	65
3.3 The Intrinsic Motivation Explanation	67
4. The Empathy-Sympathy Explanation	70
4.1 Overview	70
4.2 The Empathy-Sympathy Hypothesis	72
4.3 The Constitutive View	76
4.4 The Causal View	79
5. The Phenomenal Epistemic Explanation	81
Chapter 3: Empathy and Phenomenal Knowledge	84
1. Overview	84
2. Literature Review	89
2.1 Empathy and the Mindreading Debate	89
2.2 Empathy and the Knowledge Argument Debate	91
2.3 Supporters of the Phenomenal Epistemic Claim	94
3. Empathy as Experiential Understanding	95
4. Two Kinds of Everyday Example	103
4.1 Acquainted but phenomenally ignorant	104
4.2 Empathic epiphanies	105
5. An Argument for the Strong Phenomenal Epistemic Claim	108
6. Two Possible Objections	116
6.1 The Sufficiency of Acquaintance Claim revisited	117
6.2 Challenging the definition of empathy	120
7. Conclusion	122
Chapter 4: Empathy and Well-Being	124
1. Overview	124
2. The Prudential Epistemic Claim	125
3. The Contemporary Philosophy of Pain	132
3.1 The Badness of Pain	132
3.2 Defining Pain	133
3.3 The Phenomenal View	134
3.4 The Dislike View	137

3.5 Empathy and pain	138
3.6 The Hybrid View	139
3.7 Empathy with suffering	143
4. The Argument From Introspection	146
4.1 Recap	146
4.2 The Introspection Claim	148
4.3 The “Self-Evidence Supervenience” Claim	150
5. Conclusion	153
Chapter 5: Empathy and Normative Reasons	155
1. Overview	155
2. Reasons to Alleviate Suffering	158
3. Empathy as a Normative Guide	163
3.1 The Strong View	163
3.2 The Normative Epistemic Claim	165
3.3 The limitations of empathy	169
4. The Revised Argument From Introspection	171
4.1 Recap	171
4.2 Introspection and reasons	173
4.3 Impersonal introspection and reasons	176
4.4 The application to empathy	178
5. The Phenomenal Epistemic Explanation Revisited	180
6. Conclusion	186
Chapter 6: Empathy and Moral Deliberation	188
1. Overview	188
2. The Epistemic Defence of Empathy	189
3. Responding to the Spotlight Effect Objection	193
4. Empathy and Harm Inhibition	199
5. Empathy with Joy	201
6. Conclusion	204
Bibliography	206

Chapter 1: Introduction

1. Overview

In this chapter, I outline the main argument of this thesis and contextualize it within the contemporary philosophical debate about the role of empathy in morality. As a first step in understanding that debate it is important to clarify how the term 'empathy' has been used within it. Accordingly, I shall begin (in section 2) by explaining how contemporary philosophers have typically defined empathy to be a kind of affective matching, before classifying the ways in which they have disagreed about which kind of affective matching empathy is.

One might be inclined to assume that empathy plays a central role in morality. In section 3, I present a series of challenges to this assumption which have been posed by recent critics of empathy. The *Not Necessary Objection* holds that empathy is not necessary for morality. The *Emotional Cost Objection* says that empathy is to be avoided because it is emotionally draining for the empathizer and can lead to burnout. The *Spotlight Effect Objection* says that empathy distorts our moral deliberations by making us care too much for the person whom we are empathizing with, at the expense of those whom we are not empathizing with.

In section 4, I consider how a proponent of empathy can respond to these challenges. One general strategy is to argue for a selective use of empathy and to distinguish contexts in which empathy is helpful from contexts in which it is not. This strategy offers a middle ground in a debate that might otherwise be characterized as a choice between two extremes: on the one hand, empathizing

with as many people as possible in every situation, and on the other, abandoning empathy altogether.

This response, however, does not yet answer a crucial question: in the contexts in which empathy has an important moral role to play, what is that role? Contemporary philosophers have answered this question in three different ways. First, there is the *Epistemic Defence* of empathy, which argues that empathy is an important way of informing our moral deliberations. Secondly, there is the *Motivational Defence* of empathy, which argues that empathy is an important way of generating moral motivation. Thirdly, there is the *Relational Defence* of empathy, which argues that empathy generates a special sort of valuable interpersonal connection, one that is characterized by high levels of trust and understanding. To be developed persuasively, each of these defences must engage with the aforementioned objections to empathy.

In this thesis, my primary aim is to contribute to the Epistemic Defence of empathy. Consequently, in section 5, my next step will be to explore that defence in more detail. I argue that this defence, as it has been developed so far, faces two problems. First, it does not sufficiently explain what is special about empathy, as opposed to other ways of learning about the mental states of others, such as testimony and inference. Secondly, it is susceptible to a strong version of the Spotlight Effect Objection against empathy. Consequently, a critic of empathy can respond to the Epistemic Defence by maintaining that although empathy supplies us with information that is useful in our moral deliberations, there are other ways of getting that information that are superior to empathy in so far as they are not emotionally draining, and in so far as they do not distort our moral deliberations.

In section 6, I introduce the *Normative Epistemic Claim*, and explain how it strengthens the Epistemic Defence of empathy. The Normative Epistemic Claim says that empathy teaches us not only about which affective states other people are experiencing, but also about the reasons for action that those affective states provide us with. This claim explains what is special about empathy: empathy gives us both attributive knowledge (knowledge of what other people are feeling) and normative knowledge (knowledge of our reasons for action), whereas other ways of knowing what other people are feeling only give us attributive knowledge. The Normative Epistemic Claim also puts the defender of empathy in a much stronger position to respond to the Spotlight Effect Objection. In short, I shall argue, the Normative Epistemic Claim vindicates the role of empathy in moral deliberation. In the final section of this chapter (section 7), I explain how each subsequent chapter of this thesis contributes to my defence of the Normative Epistemic Claim.

2. Empathy

2.1 Affective Matching

Generally speaking, contemporary philosophers have agreed that empathy is a form of affective matching.¹ That is, they say that empathizing with another person involves feeling an affective state (an emotion, mood, or physical pleasure

¹ See, for example, Darwall (1998), Slote (2010), Coplan (2011), Prinz (2011a and 2011b), Maibom (2014), and Song (2015). For an influential example of this definition being applied in biology and psychology, see Sober and Wilson (1998), and Eisenberg (2011). Nussbaum (2001) and Goldie (2002) offer a different definition, on which empathy involves imaginatively reconstructing another person's perspective and may or may not involve actually feeling what the other person feels. For a rival view of empathy (which focuses on recreating thoughts, rather than affects), see Stueber (2006).

or pain) that is similar to an affective state that the other person is feeling. To put things more formally, they say that there is an *Affective Match Condition* on empathy.

The Affective Match Condition: For it to be the case that person A is empathizing with person B with respect to B's feeling S (an affective state), it must be that A feels S*, where S* is similar to S.

By way of illustration, empathizing with an angry person will involve feeling anger and empathizing with a distressed person will involve feeling distress. On this approach, the term 'empathy' does not refer to a specific emotion, but rather to a state of mirroring the emotion of another person (whatever that emotion is).

Different philosophers have then developed their account of empathy in different ways. First, some offer clarity on how close the affective match needs to be, in order for it to qualify as a case of empathy. In other words, they specify a requirement for how similar S and S* need to be.² Secondly, a philosopher may place a requirement on who S* is felt for. On Maibom's account of empathy, for example, S* must be felt for the other person (that is, the person whom A is empathizing with). To empathize with a distressed friend, on Maibom's account, I must feel distressed for them.³ This allows one to make a clear distinction

² A 'relaxed view' is that the two affective states only need to be of the same valence (i.e. either both positive, or both negative). My view, which I state in chapter 3, is that the affective states need to be of the same type (e.g. both fear, or both anger, or both sadness), and of approximately the same intensity. Where the affective states are emotions, there is a further question about whether they need to have the same objects. My view is that they do not need to.

³ It should be noted that, in writing this thesis, I have not applied a single rule for deciding which pronouns to use when discussing examples. I have sometimes used 'they' and have also sometimes use gender-specific pronouns.

between empathy and personal distress, for in the latter case I would feel distressed for myself.⁴

Thirdly, some philosophers add a knowledge condition on empathy, which requires that the empathizer not only match the affective state of the other person, but also that they know which affective state the other person is feeling. In other words, the requirement would be that the empathizer knows that the other person is feeling S. For example (according to this requirement), for me to empathize with an irritated friend, I must not only feel a similar irritation, but also know that my friend is feeling irritated. I shall call this knowledge (of which affective state another person is feeling) *attributive knowledge* because it involves attributing an affective state to another person.⁵

On some accounts of empathy, the affective match is framed as a cause of the attributive knowledge. It has been argued that replicating the affective states of others is an important means through which we come to know what other people are feeling (Goldman, 2006). An intuitive explanation for this runs as follows: if we have matched another person's affective state, and we know both what we are feeling and that what we are feeling is a match of the other person's affective state, then we can infer what the other person is feeling.⁶

⁴ One might think that this is a further elaboration of the similarity condition, in so far as my friend is also (presumably) feeling distressed for themselves. If I empathize with my friend, then (on Maibom's account of empathy) both my friend and I will feel distressed for the same person (my friend). However, on a different interpretation of what similarity involves, my emotional state could be said to be more similar to my friend's emotional state when I feel distressed for myself. In that case, both my friend and I would feel distressed for ourselves.

⁵ One might also refer to this as 'cognitive empathy' (Eslinger, 1998) or as 'empathic accuracy' (Ickes, 1993).

⁶ It should be noted that on Hume's model of empathy, it is the attributive knowledge that gives rise to the affective match (and not vice versa). On Hume's account, we can form an idea of another person's affective state that is so lively that it becomes a similar affective state. I discuss Hume's view further in the section on affective contagion below.

Fourthly, philosophers have tended to add a causal condition to their definition of empathy, which specifies that the empathic affect (S^*) must arise through a particular causal mechanism. For example, Coplan (2011) requires that the empathic affect must arise through perspective-taking. Perspective-taking is a consciously mediated process in which the empathizer imagines being in the position of the person that they are empathizing with, and thereby comes to feel something similar to what that person is feeling. Slote (2010) ties empathy to the mechanisms of affective contagion. Mechanisms of affective contagion are a set of processes that are not consciously mediated, through which the empathizer comes to feel S^* by either perceiving the affective expressions of the person that they are empathizing with (e.g. their tears of sadness), or by perceiving the immediate cause of the affective state in the other person (e.g. seeing them open a letter that is known to contain bad news). A hybrid approach (Goldman, 2006; Prinz, 2011a) is to require that the affective match arises through either affective contagion or through perspective-taking.⁷ I shall say more about perspective-taking and affective contagion below.

Philosophers who define empathy to be a kind of affective matching frequently go on to distinguish it from sympathy.⁸ Those philosophers define sympathy as follows.

⁷ It should be noted that Goldman uses different terminology to that which I employ here: roughly speaking, he refers to affective contagion as 'low level simulation' and to perspective-taking as 'high level simulation'. For articulations of the hybrid approach within psychology, see Eisenberg and Strayer (1987) and Hoffman (2000).

⁸ See Darwall (1998), Slote (2011) and Maibom (2014). For similar definitions in biology and psychology, see Sober and Wilson (1998) and Eisenberg (2011). Nussbaum (2001) uses the term 'compassion' to refer to what I am here calling 'sympathy'.

Definition of Sympathy: Person A is sympathizing with person B just if A feels sad, concerned or distressed for B because A perceives or believes that B is in need.

There are then three kinds of case in which the notions of sympathy and empathy can come apart. First, empathy can involve affective states like guilt, rage, boredom, depression and physical pains whereas sympathy cannot. It is thus one thing to empathize with a friend's depression, and another thing to sympathize with it. Second, one might allow (though, as noted above, Maibom would not) that one can empathize with a sad person by feeling sad for oneself (thus mirroring the way that the other person feels sad for themselves), but this would not count as a case of sympathy because in sympathy one must feel sad for the other person. Finally, empathy and sympathy come apart in cases where a person is in need, but is not themselves feeling any kind of negative affective state. Sympathy for such a person would involve feeling sad, concerned or distressed for them, but empathy with them would either be impossible (if they are not feeling *any* affective state) or would involve feeling a positive affective state (in the case that they, not realizing their plight, are feeling such an affective state). Using the term 'concern' to refer to what I am here calling 'sympathy', Jesse Prinz gives an example of the latter scenario:

When we see a drug addict take another hit, she may exhibit a euphoric response. Empathy might induce joy in this case, but concern makes us worry about the addict's well-being. (Prinz, 2011b: p. 230)

Of course, even if there are cases where empathy and sympathy diverge, there are also cases that count as both empathy and sympathy. If I imagine being my friend who is feeling sad because his dog has just died, and thus come to feel sad for him, this may count both as a case of empathy and as a case of sympathy (in

such a case we could not say that empathy caused the sympathy, but should say that the perspective-taking caused the sympathy). I shall discuss the relationship between empathy and sympathy further in the next chapter.

It is worth flagging a possible source of confusion about the terminology that I am using here. The 18th century writings of David Hume and Adam Smith are generally regarded to be the first expositions of empathy in Western thought. However, when they are discussing (what we would call) empathy, they refer to it using the term 'sympathy'. The term 'empathy' was not introduced into the English language until the early 20th century.⁹ 'Sympathy' appeared an apt choice of word for Hume and Smith because, according to its Greek etymology, it literally means 'together in feeling' (sun-pathos). In the years since then, however, the meaning of sympathy has migrated into its contemporary usage, creating a space for a new term (empathy) to enter (Depew, 2005).

Having summarized how the term empathy has been used in contemporary moral philosophy, I shall shortly (in section 3), offer a summary of the recent criticisms that have been made of empathy. This summary will serve as a useful backdrop against which to explain how and why I defend the moral significance of empathy.

Before proceeding to these tasks, however, I spend the next two subsections going into more detail about the two kinds of matching mechanism

⁹ The first usage of the term 'empathy' is typically attributed to Titchener (1909). Titchener used that term as a translation of the German term 'Einfuehlung'. After emerging as a technical term in aesthetics in the late 19th century, 'Einfuehlung' had become important in early 20th century German psychology. Although Titchener used a term with Greek roots (em-pathos, literally meaning feeling-into) for his translation, he was not drawing on a Greek tradition of using that term in the way that he used it. For a detailed history of the term 'empathy', see Depew (2005). Coincidentally, Titchener grew up in Chichester, a small city in the South of England, which is also where I am from.

identified above: perspective-taking and affective contagion. For each kind of mechanism, I clarify the different forms that it may take, discuss the conditions under which it gives rise to an affective match, and consider whether it might also give rise to attributive knowledge. I begin with perspective-taking.

2.2 Perspective-Taking

Above, I defined perspective-taking as a consciously mediated process of imagining being in the situation of another person, and thereby coming to feel emotions that are appropriate to that situation. The emotional force of perspective-taking derives from the power of our imagination: our emotions do not just respond to the situations that we believe ourselves to be in, but also to the situations that we imagine ourselves to be in. For example, I might be feeling calm and relieved after meeting a deadline, but by imagining myself to be in the situation of a friend who has upcoming exams, I can bring myself to feel stressed and anxious.

Following Stotland (1969), it is common to distinguish between two forms of perspective-taking. In self-oriented perspective-taking, I imagine being myself in the situation of the other person. In other-oriented perspective-taking, I imagine being the other person in their situation. These two processes produce different results when I respond to the other person's situation in a way that is different from the way in which they respond to it (Hygge, 1976; Jarymowicz, 1992). In the above example, I was using self-oriented perspective-taking to imagine how I would feel in my friend's situation. Yet my friend and I might respond differently to the pressure of exams. Perhaps my friend feels excited about his upcoming

exams. Other-oriented perspective-taking would require me to adjust for the differences between my friend and me to imagine not how I would feel in his situation, but how he feels in his situation.

Perspective-taking can give rise to an affective match, although that may not always be its intention. In Adam Smith's account of sympathy, he claims that we use self-oriented perspective taking (which he calls sympathy) to compare how another person is feeling with how we would feel in their situation (Smith, 1759/1975). This comparison is then used to evaluate the other person's feeling. Roughly speaking, if we would react in the same way that the other person is reacting, we are likely to approve of their reaction.¹⁰ On this approach, the purpose of perspective-taking is not to produce an affective match.

If we are using self-oriented perspective-taking to produce an affective match, then three conditions will need to be met.¹¹ First, our reactive dispositions must be sufficiently similar to those of the other person. If they are not, we will need to use other-oriented perspective-taking. Secondly, we must have sufficient information about the other person's situation. If we are missing information about the situation, we may respond in a way that is different to the way that we would respond if we had all of the relevant information. For example, if I did not know that the friend with upcoming exams had already done so well on previous exams

¹⁰ See, for example: "passions of human nature, seem proper and are approved of, when the heart of every impartial spectator entirely sympathizes with them, when every indifferent by-stander entirely enters into, and goes along with them" (Smith, 1759/1975: p. 81). For a more extended discussion of Smith's view, which explores what it means to be an 'impartial spectator', see Kauppinen (2014).

¹¹ It should be noted that even if self-oriented perspective-taking gives rise to an affective match, it may fail to meet the other conditions for empathy. For example, it has been suggested that self-oriented perspective-taking can lead to personal distress, rather than to empathy (Batson, Early & Salvarini, 1997). If you imagine you had you would feel in another person's situation, you can become pre-occupied with worries about how to avoid that situation, and lose sight of the plight of the other person.

that the upcoming exams could have no impact on his grade, then in imagining his situation I would come to feel stressed, whereas he felt relaxed. On the other hand, we must also quarantine information that we have but that the other person lacks (Van Boven & Loewenstein, 2003). If I know that the friend's exams have doubled in difficulty, but he does not, then imagining being in his situation may lead me to feel more stressed than he actually feels.

If I have a different reactive disposition to the person that I am empathizing with, then I will need to use other-oriented perspective-taking in order to achieve an affective match. In this case, I will not only need to know about the other person's situation (and what they know about it), but I will also need to know about the reactive dispositions of the other person, so that I can adjust my perspective-taking accordingly.¹² For example, if I know that my friend finds delight in taking exams, but I detest them, then I could try to take up his perspective by imagining doing something that I love doing, such as visiting a theme park.

As a final point in this section, it is important to note that where perspective-taking does give rise to an affective match, it can also give rise to attributive knowledge.¹³ For example, in self-oriented perspective-taking, we might reason as follows: 'if I was in the other person's situation, I know through perspective-taking that I'd feel this affective state. My reactive dispositions are

¹² In other-oriented perspective-taking, it is also important that we are able to mimic the reactive dispositions of the other person. For an interesting discussion of the challenge involved in mimicking dispositions that we may regard as evil, see Morton (2011).

¹³ For an interesting discussion of perspective-taking as a route to attributive knowledge, see Goldman (2006) and Matravets (2011). It should be noted that Goldman uses the term 'high-level simulation' to refer to what I am here calling 'perspective-taking'. However, an important distinction is that Goldman does not require that high-level simulation is consciously initiated. Accordingly, the term 'high-level simulation' covers both cases of consciously initiated perspective-taking (which are, according to my definition the only cases of perspective-taking) and non-consciously initiated perspective-taking (which I classify as a kind of affective contagion).

similar to those of the other person, and thus I can infer that he also feels this affective state.'

2.3 Affective Contagion

A second kind of mechanism that can lead to affective matching is what I shall refer to as affective contagion. Affective contagion is the spread of affective states from one person to another through a contagion mechanism. I take a contagion mechanism to be a psychological mechanism through which one person, A, comes to feel an affective state that another person, B, is feeling, where this mechanism (i) is not consciously initiated, and (ii) requires that person A has either perceived the affective expressions of B (e.g. seen B cry and look sad) or perceived B in the situation which immediately gave rise to B's affective state (e.g. B opening a letter that is known to contain bad news).

David Hume is often regarded as the first philosopher to articulate the phenomenon of affective contagion.¹⁴ His notion of sympathy is closely related to affective contagion, and he describes it as follows:

When any affection is infus'd by sympathy, it is at first known only by its effects, and by those external signs in the countenance and conversation, which convey an idea of it. This idea is presently converted into an impression, and acquires a degree of force and vivacity, as to become the very passion itself, and produce an equal emotion, as any original affection. (Hume, 1739/1978: p. 317)

Hume thus offers us a three-step process of affective contagion: (i) we notice the outward expression of another person's affective state, then (ii) we form an 'idea'

¹⁴ For a useful overview of Hume's philosophy of sympathy, see Nillson (2003) and Ilyes (2017). It should be noted that it is probable that Hume's account of sympathy was inspired by his reading of Spinoza (Cassidy, 1979).

of their affective state, and then (iii) this idea transforms into an affective state that we feel. Elsewhere, Hume suggests that whether or not this third step happens depends on our relationship to the person in question. Roughly speaking, the more similar we are to them, and the closer we are to them spatio-temporally, the stronger the affective contagion will be.

It is important to note that on Hume's account of affective contagion, attributive knowledge (what he calls our idea of the other person's affection) precedes the affective match. For Hume, we first know what another person is feeling (we form an idea of their affection), and then come to feel what they are feeling. Contemporary approaches to affective contagion have typically inverted this relationship. It has been argued that it is through affective contagion that we come to know what other people are feeling (D'Arms, 2000; Goldman, 2006). If affective contagion precedes attributive knowledge, then the contagion must be arising through an alternative, non-Humean, mechanism.

One influential account of affective contagion has been offered by the research into mimicry. We naturally mimic the physical expressions (including facial expressions, gestures, and actions) of other people (McIntosh et al., 1994). For example, if you show a respondent a series of faces that express either anger or happiness, there will be a measurable difference in the muscular activity of the respondent's face, depending on which kind of face they are observing (Dimberg, 1982, 1988). If a respondent watches an arm wrestle, they will mimic both the facial expressions and the arm movements that they observe (Berger and Hadley, 1975).

When we mimic the physical expressions of another person, it tends to induce a similar affective state in us to that which caused the expression in the other person (Adelman and Zajonc, 1989). For example, if we are talking with a person who is happy, and thus smiling, we will tend to mimic their smile, and mimicking their smile will make us also feel happy. Subjects who intentionally amplify their mimicry of another person's facial expressions experience heightened mimicry of the other person's affective states (Vaughan and Lanzetta, 1981).¹⁵

More recently, there has been considerable excitement about the discovery of mirror neurons (Iacoboni, 2008; De Waal, 2009; Singer, 2009). The term 'mirror neurons' was first used in a paper describing research that had been done on the brains of macaque monkeys (Gallese et al., 1996) and referred to neurons that discharged both when the monkey was performing a particular action (grasping, holding or tearing) and when the monkey observed another monkey performing that action. One way of framing this is to think of there being contagion of motor impulses. Mimicked motor impulses may or may not lead to mimicked behaviour, depending upon whether the impulses are inhibited.¹⁶

Since those early experiments, there have been many studies of mirror neurons in the human brain (a useful summary can be found in Bernhardt and Singer, 2012). The term 'mirror neuron' has come to be used for any neurons that have the general property of mirroring, which is the property of firing both when

¹⁵ The link between facial expression and emotion has also been shown in studies that do not involve mimicry of physical expressions (Ekman, 1992). One such study found that asking subjects to say 'e' led to an increase in their happiness. There may something to saying 'cheese' after all!

¹⁶ Compulsive behavioural imitation has been shown in patients with brain damage that impedes inhibition of motor impulses (Lhermitte et al, 1986).

the subject is in a mental state and when they observe another subject either expressing that mental state, or being in a situation that stimulates that mental state.

Mirror neurons have been shown to operate in many different parts of the brain, and to provide means of contagion for many different types of mental state. By way of brief survey, there are mirror neurons that fire: when you observe another person being touched (Keysers et al., 2004), when you observe the stimulation of pain in another person (Singer et al., 2004), when you see the facial expressions of pain in others (Botvinick et al., 2005), when you observe other people react with disgust (Wicker et al., 2003), when you observe anxiety in others (Prehn-Kristensen et al., 2009) and when you see someone being socially excluded (Masten et al., 2011). Mirror neurons have also been found to fire when we observe positive emotions in others (Jabbi et al., 2007).

What is the relationship between mirror neurons and the mimicry mechanism discussed above? On the one hand, it may be that mirror neurons explain the behavioural mimicry component of that mechanism. On the other, they may provide an alternative mechanism (or even mechanisms). For example, in the Singer et al. study cited above, participants came to feel pain in their hand after watching a syringe being inserted into the hand of another person. Participants did not perceive the behavioural reaction of the other person, and thus the affective match cannot be explained through mimicry of that reaction. Instead, the affective match arose because the participant perceived the situation that had stimulated the affect in the other person.

In this kind of case, what we seem to have is an automated version of the process of perspective-taking that I described above. The empathizer seems to imagine themselves being in position of the other person (that is, having a syringe inserted in their hand), and thus comes to feel what they would feel in that situation. In his discussion of sympathy, Adam Smith gives several everyday examples of this phenomenon of automated perspective-taking:

When we see a stroke aimed and just ready to fall upon the leg or arm of another person, we naturally shrink and draw back our own leg or our own arm; and when it does fall, we feel it in some measure, and are hurt by it as well as the sufferer.... Persons of delicate fibres and a weak constitution of body complain, that in looking on the sores and ulcers which are exposed by beggars in the streets, they are apt to feel an itching or uneasy sensation in the correspondent part of their own bodies. (Smith, 1759/1975: p. 10)

In these cases, we are not responding to the expression of the affect by the other person. As Smith describes these cases, we do not see the other person shrink and draw back their leg or arm, and nor do we see the beggars complain or scratch themselves. Rather, the affective contagion arises because we perceive the other person's situation and imagine ourselves being in it. As I discussed in the section on consciously mediated perspective-taking above, this process will yield an affective match only when we react in a way that is similar to the way in which the other person reacts.¹⁷

It is important to distinguish between conscious and automatic perspective-taking. I reserve the term 'perspective-taking' for the former, and consider the latter to be a type of affective contagion. This distinction is important because if the perspective-taking is automated, then we may not be aware that we are doing

¹⁷ If the person who is being beaten feels no fear, or if the beggar is desensitized to their sores, then automated perspective-taking may fail to produce an affective match with them.

it, and we may not be aware of the affect that it generates (or of how it came about). In chapter 3, I argue that this kind of awareness is required in order for cases of affective matching to count as cases of empathy. Accordingly, automatic perspective-taking, even if it leads to an affective match, is less likely to give rise to empathy than its consciously mediated counterpart.

Where affective contagion successfully gives rise to an affective match, it can also give rise to attributive knowledge. Indeed, where affective contagion is prompted by perceiving the affective expressions of other people, it has been argued that it is more epistemically reliable than perspective-taking (D'Arms, 2000). One worry here is that although affective contagion might lead to a match of emotional states, it would not ensure that the two emotional states had the same object. For example, through mimicry, I might match the delight of a friend who had just received a promotion. My delight, however, would not be about a promotion, and I may remain entirely ignorant of why my friend was feeling delighted.

Having said something about what empathy is, and the mechanisms through which it can arise, I now move to review the recent criticisms that have been made of empathy. It should be noted that these criticisms tend to assume that empathizing with other people is something that we can control, such that we can choose to do more or less of it. This assumption is, by definition, true with regard to empathy that arises through perspective-taking, but is more problematic with respect to empathy that arises through affective contagion. I have said that such contagion is automatic. However, we can vary our exposure to situations that are likely to stimulate affective contagion (e.g. do you pay attention to a

homeless person as you walk past them?), and we may also be able to train ourselves to feel more or less affective contagion in any given situation.

3. Against Empathy

3.1 *Bucking the trend*

In recent years, there has been a significant push-back against the idea that empathy is an unqualified moral good. Most recently, psychologist Paul Bloom published a book, *Against Empathy*, in which he forcefully argues that we would be better off if we stopped empathizing with others (Bloom, 2016).¹⁸ Bloom's objections are similar to (and make reference to) earlier objections developed by Jesse Prinz (Prinz, 2011a and 2011b).

Both Bloom and Prinz take themselves to be writing within an intellectual context in which empathy is widely assumed to be a good thing. As Bloom reports:

I've learned that being against empathy is like being against kittens – a view considered so outlandish that it can't be serious. (Bloom, 2016: p. 15)

In the introduction to his article *Against Empathy* (Prinz, 2011b), Prinz airs a similar view:

Would it be good to have empathetic juries and judges? What about empathetic activists or hospital ethicists? Should we pay special attention to the moral theories of empathetic philosophers? Intuitively, the answer might be "yes." Empathy is a thick concept, and it connotes praise. (ibid.: p. 214)

¹⁸ Bloom's book extends arguments that he had developed in earlier magazine articles (Bloom, 2013 and Bloom, 2014).

Prinz then continues in a way that suggests that, in his experience, this intuition in favour of empathy is not typically argued for:

But an endorsement of empathy requires more than a warm fuzzy feeling. We need an argument for why empathy is valuable in the moral domain. (ibid.: p. 214)

My anecdotal experience from talking to friends and colleagues has been that the popularity of empathy is indeed widespread. As further evidence for this, one can also point to the long and growing list of books published in the last twenty years which offer an enthusiastic defence of empathy to a general audience (Gordon, 2005; De Waal, 2009; Trout, 2009; Baron-Cohen, 2011; Krznaric, 2014; Bazalgette, 2017).

3.2 Morality without empathy

Standing in opposition to this trend, both Prinz and Bloom think that one can be a moral person without empathizing with others. On the face of it, we might find this claim to be surprising. It is important to remember, however, that by the term 'empathy', they mean, as outlined above, a form of affective matching. To empathize with another person is, on this approach, to feel what they feel. However, as Bloom notes, the everyday usage of the term 'empathy' tends to be much broader.¹⁹ The term 'empathy' can be used to refer 'to everything good ...as a synonym for morality and kindness and compassion' (Bloom, 2017: p. 3). Prinz and Bloom are not against empathy in that (broad) sense. Rather, they are against empathy as affective matching.

¹⁹ For an analysis of the everyday usage of empathy, see also Battaly (2011).

It will be recalled that in the section above I made distinctions between empathy and sympathy, and between empathy and attributive knowledge. Prinz exhibits a favourable attitude towards sympathy (which he calls ‘concern’), whilst being against empathy (Prinz, 2011b: p. 230-231). Bloom acknowledges that attributive knowledge is extremely useful in sensitively navigating our social lives, but maintains (correctly) that attributive knowledge is not empathy (Bloom, 2017: p. 3). While it may be true that empathy is one way of achieving attributive knowledge, Bloom can maintain that there are other ways of achieving such knowledge that we should use instead.

In envisioning a morality without empathy, Prinz takes aim at a specific historical target. He focuses on critiquing the view that empathy is necessary for making moral judgements (Prinz, 2011b: pp. 216-224).²⁰ He attributes this view to David Hume, and notes that a contemporary version of Hume’s view has been defended by Michael Slote (2010). Prinz believes that Hume and Slote are wrong. For Prinz, we can make moral judgements without the use of empathy.

By way of a brief summary, Prinz’s general strategy here is to identify examples in which we form moral judgements about wrongdoing, but in which it is not plausible to suggest that we make these judgements by empathizing with a person who has been harmed by that wrongdoing. First, there are cases in which we ourselves are the person who has been wronged. For example, if I have had something stolen from me, I may judge that I have been wronged, but it seems

²⁰ There are two ways in which empathy could be necessary for moral judgement: (i) empathy could be synchronically necessary for moral judgement, such that empathy is required on each and every occasion that one makes a moral judgement or (ii) empathy could be developmentally necessary for moral judgement, such that the exercise of empathy is a necessary part of the moral education that gives an agent the capacity to make moral judgements as an adult (even if those moral judgements can then be made without the synchronic use of empathy). Prinz argues against both claims. For further discussion, see Denham (2017).

implausible to suggest that I needed to empathize with anybody in order to make this judgement. Secondly, there are acts that we judge to be wrong, but which have not harmed anybody. Prinz cites two examples: consensual incest between two adult siblings, and the destruction of an unpopulated part of the environment. Thirdly, there are cases (such as tax evasion) in which, although people might be harmed by an act of wrongdoing, the harm is spread amongst a very large number of individuals, and is thus unlikely to be something that we attend to in forming our moral judgement.

Prinz also considers, and rejects, the idea that empathy is necessary for moral motivation (Prinz, 2011b: pp. 225-227). He claims that, at best, empathy for a suffering person only generates weak motivation to help them.²¹ On Prinz's approach, there are alternative, more powerful, sources of moral motivation, such as guilt and anger in response to moral transgressions. Prinz is also open to the idea that sympathy (which he calls 'concern') is an important moral motivator (ibid.: p. 230-231).

3.3 The Emotional Cost Objection

According to the critics of empathy, it is not just that empathy is not necessary for morality, but that where we do use empathy, it is problematic. The first problem associated with empathy is that it is emotionally costly for the empathizer (Bloom, 2016: p. 136-45). Empathy with suffering is painful and

²¹ In the next chapter, I summarize evidence that empathy is, contra Prinz, a powerful source of moral motivation. Prinz's view that empathy is at best only a weak motivator is perhaps inconsistent with his worries, discussed below, that empathy is a source of bias. Moreover, Prinz does not discuss the evidence that empathy is a source of sympathy, which he considers to be a powerful motivator. For further critique of Prinz's argument, see Denham (2017). I agree with Prinz, however, that empathy is not necessary for moral motivation.

draining, and can have negative health effects for the empathizer (Singer and Klimecki, 2014).²² It may even be debilitating, to the extent that the empathy makes the empathizer less able to help the person that they are empathizing with. By definition, when you empathize with a person who is feeling a painful affective state (e.g. feeling anxious) you will feel a similar, and similarly painful, affective state (e.g. you will feel anxious).²³ We might all be familiar with the strain of empathizing with a friend who is going through a difficult period.

If a person has a job in which they repeatedly empathize with the suffering of others, then their empathy for others may impede their ability to do that job. This will particularly be the case if the suffering that they are empathizing with is severe. For this reason, there are worries raised about empathy burnout in professions such as healthcare and psychotherapy (Hoffman, 2002). Moreover, given the prevalence of modern news stories involving extreme suffering, it might be suggested that anyone who responded to such stories with empathy would also be at the risk of burnout (Kaplan, 2011).

None of this, of course, entails that we should not empathize with those who are suffering. Rather, it says only that there is a cost to doing so. Even if over-empathizing is to be avoided, there might still be important benefits to the selective use of empathy which make it worth the emotional cost that it entails.

²² As an alternative to empathy, Singer suggests that we train ourselves to feel compassion in response to the plight of others. Singer's notion of compassion is to be distinguished from sympathy. Whilst the latter is a negative feeling, compassion is a warm and pleasant emotional state. According to Singer, compassion is effective in motivating action and is not draining in the same way that empathy is. Singer's approach was inspired in part by conversations with the Buddhist monk Matthieu Ricard.

²³ Of course, if you empathize with a joyful person, your empathy will involve feeling joyful affective states, which might be uplifting and empowering. In general, I think that the significance of empathy with joy has been neglected within the recent debates. I shall return to consider its importance at the end of this thesis.

3.4 The Spotlight Effect Objection

A second line of criticism against empathy is that it is a poor moral guide. This criticism is raised extensively by both Bloom (Bloom, 2016: ch. 1) and Prinz (2011b: pp. 227-230). At the heart of this line of criticism is what I call the *Spotlight Effect Objection*, which I shall now summarize.

The Spotlight Effect Objection allows that empathy may often be used to help form our moral judgements (even if it is not necessary for making such judgements). It says, however, that where it is used to shape our moral judgements, it can make them worse. It does this by making us think that we should help the person that we are empathizing with, regardless of whether this is in fact true. Empathy looks like it has a favourable impact on our decision-making when we are considering a choice between helping an individual in need, and doing nothing. If empathy makes us think we should help that person, then this may well be a good thing. The problem arises in cases where there are multiple people whose interests are at stake. Empathizing with one person will then, the objection says, lead us to prioritize that person's needs and to fail to give sufficient weight to the needs of other people.

There are a variety of contexts in which this might happen. An empathic judge or jury might, depending on whether they empathize with the victim or the accused, be excessively harsh or lenient towards the accused. Someone allocating medical resources may treat a lower-priority patient first simply because they happened to empathize with them. A donor might give money to an inefficient charitable project because they empathized with someone who would be helped by that project.

Below, I shall return to consider the Spotlight Effect Objection in more detail, and will discuss how to respond to it. For now, it is important to note that it is closely related to another objection that is often made against empathy.²⁴ The *Partiality Objection* says that there is a problematic bias in the pattern of who we tend to empathize with: put simply, we tend to empathize more with people who are in our social group (Stürmer et al., 2005; Brown et al., 2006). When combined with the Spotlight Effect Objection, it implies that our habitual usage of empathy will incline us towards thinking that we should prioritize the interests of those people, at the expense of others. In other words, it implies that our habitual usage of empathy will make us partial towards our in-group in a problematic way.

One might respond to this objection by arguing that such empathy-induced partiality is morally appropriate (Slote, 2010). That is not the strategy that I shall adopt here. Instead, I seek to side-step this objection by framing it as a criticism of our habitual patterns of empathy, and not as a criticism of empathy itself. As later examples will demonstrate, it is possible to counter biases in our habitual use of empathy by intentionally cultivating empathy towards people in different social groups.

4. Defending Empathy

4.1 Selective Empathy

If one takes the Spotlight Effect Objection seriously, one must acknowledge that there are situations in which empathy can lead our moral

²⁴ See, for example, Prinz (2011a, pp. 226-228).

decision-making off course. A promising response to this objection is to maintain that although there are such situations, there are also situations in which it remains morally appropriate to empathize with others. The task is then to delineate when and where we should use empathy (Carse, 2005; Song, 2015). For example, a defender of empathy could concede to Prinz and Bloom that it is not appropriate for decision-makers to rely on empathy in contexts such as legal trials, the allocation of public resources, and the direction of philanthropic efforts. They could maintain, however, that empathy remains helpful, even vital, in other contexts.

One context where empathy may be of particular value is in personal relationships. Along this line, the use of empathy has been defended in psychotherapy (Rogers, 1980), social work (Howe, 2013), conflict resolution (Rosenberg, 2003), and medical care (Halpern, 2001). Prinz himself seems open to this approach. At the start of *Against Empathy*, he writes:

We would all like to have empathetic friends. It is nice when your “near and dear” can pick up on your moods, revel in your achievements, and mourn your losses. (Prinz, 2011b: p. 214).²⁵

However, arguments have also been made for the importance of empathy outside of the realm of our personal connections. Most notably, it has been argued that empathy is a powerful way of mobilizing appropriate levels of concern for oppressed and / or underprivileged social groups (Hoffman, 2011 and 2013). For example, Kznaric claims that empathy for the suffering of slaves played a key role in 19th century abolitionist movements (Kznaric, 2014). Even if we habitually

²⁵ Bloom steers us away from using empathy, even in personal relationships. In large part, this is because of the Emotional Cost Objection described above (Bloom, 2016: ch. 4).

tend to empathize with those in our social group, we can learn to redirect our empathy towards out-group members, with powerful effects.

Whilst this 'selective empathy' response is appealing in so far as it highlights contexts in which empathy can be valuable, it has not yet explained the way in which empathy is supposed to be valuable in those contexts. There are three ways that contemporary philosophers have sought to defend the value of empathy: the *Epistemic Defence*, the *Motivational Defence*, and the *Relational Defence*. I review each briefly here before moving to explain how I intend to contribute to the Epistemic Defence of empathy.

4.2 The Epistemic Defence

The Epistemic Defence says that empathy is a way of finding out about the affective states of other people, and that such insight is helpful in the process of deliberating about how we are to act (Masto, 2015; Song, 2015). For example, in order to respond sensitively to the suffering of others, we need to know when they are suffering, and we need to know which of our actions would help alleviate their suffering. Similarly, in order to avoid harming others, we need to know which of our actions would cause them pain. This kind of information is important in our personal connections (e.g. in responding to a friend who is going through a difficult time), but it could also be important in the allocation of public resources (e.g. in designing legal systems that deal sensitively with victims) and in the direction of philanthropic efforts (e.g. in evaluating whom we should help first).

A key part of this defence is the claim that empathy is a way of learning about the affective states of others. To use the terminology that I introduced

above, this claim says that empathy is a way of acquiring attributive knowledge about the affective states of other people. This claim has been defended within the recent debates about our folk psychological ‘mind-reading’ abilities (Goldman, 2006 and 2011).²⁶ However, it is important to note that even if empathy is one important route through which we come to acquire attributive knowledge, it is not the only route. We can also learn about what other people are feeling through testimony (e.g. through asking them how they are feeling) or through making inferences from what we know about their situation, or from their vocal and behavioural expressions.

The Epistemic Defence of empathy does not say that empathy is necessary for making moral judgements about how we should act. Nor does it say that empathy is sufficient for making a correct moral judgement about how we should act. Rather, it says that empathy is one way of helpfully informing our moral judgements about how we should act. Nevertheless, the Epistemic Defence must still grapple with the Spotlight Effect Objection against empathy: if empathy can consistently lead us astray in our moral deliberations, how can one say that it is a useful epistemic guide? Below, I shall give further consideration to the Spotlight Effect Objection, and will discuss how a defender of empathy can respond to it.

²⁶ As far as I know, that literature does not consider the implications of this claim for moral deliberation.

4.3 The Motivational Defence

The Motivational Defence of empathy holds that empathy is an important source of moral motivation (Darwall, 1998; Nussbaum, 2001; Maibom, 2006; Simmons, 2014). Evidence for this argument has been found in the empirical research of Daniel Batson, who holds that empathy is an important source of altruistic motivation for helping those in need (Batson, 1991 and 2011).²⁷ It should be noted, however, that Batson does not argue that empathy is necessary for moral motivation. Rather, Batson claims only that empathy is sufficient for altruistic motivation, and that altruistic motivation is one form of moral motivation. I shall discuss Batson's work, and the relationship between empathy and altruism, at greater length in the next chapter.

The Motivational Defence of empathy is typically framed in terms of how we might use empathy to respond to the suffering of strangers, but it could also be applied to how we care for our friends and family. Even in our close personal connections, we may benefit from the greater degree of altruistic motivation that empathy can bring.

One might think that the Epistemic Defence and the Motivational Defence are closely connected. The thought may be: if you know about someone else's suffering, you will be motivated to help them. However, these two approaches to defending empathy have typically been developed in isolation from one another. On the one hand, philosophers who have offered the Motivational Defence of

²⁷ As I discuss in the next chapter, Batson uses the term 'empathy' to refer to what I am calling 'sympathy'. Nevertheless, his research provides evidence that empathy (as affective matching) is a source of altruistic motivation. For further evidence that empathy is a source of pro-social motivation, see Eisenberg and Miller (1987) and Hoffman (2000). Although Batson focuses on the role of empathy in motivating helping behaviour, there is also an extensive psychological literature discussing the role of empathy in harm inhibition (for an overview, see Maibom, 2014: pp. 23-27).

empathy have explained the motivating power of empathy without referring to its epistemic function.²⁸ On the other hand, philosophers who have offered the Epistemic Defence of empathy have not considered its motivational implications.

It is important for a proponent of the Motivational Defence to acknowledge the Spotlight Effect Objection that I outlined above. In other words, the proponent of the Motivational Defence should acknowledge that empathy has a downside in so far as it can lead our behaviour astray. However, one can acknowledge this whilst maintaining that there are contexts in which empathy-induced motivation is morally appropriate. Indeed, the two psychologists who are most strongly in favour of empathy both defend it despite acknowledging at length the ways in which it can lead to immoral behaviour (Hoffman, 2000: ch. 8; Batson, 2011: ch. 8).

4.4 The Relational Defence

The Relational Defence of empathy holds that empathy is a valuable component of personal relationships, independently of the way that it informs and motivates moral behaviour (Carse, 2005; Song, 2015). On this view, when you empathize with a friend who is going through a difficult time, you develop a deeper, and more valuable, form of connection with that friend. Regardless of whether your empathy for that friend informs or motivates further helping, the Relational Defence says the empathy that you have for your friend is itself of value, both for you and for them.

²⁸ I shall discuss competing explanations for the motivating power of empathy in the next chapter.

One context in which the Relational Defence has been developed is in the literature discussing how to conduct counselling and psychotherapy successfully (Clark, 2007). In that literature, a close connection has been drawn between empathizing with a person and listening to them deeply. Carl Rogers, one of the leading advocates for the use of empathic listening in psychotherapy, describes its benefits as follows:

When a person realizes he has been deeply heard, his eyes moisten. I think in some real sense he is weeping for joy. It is as though he were saying, "Thank God, somebody heard me. Someone knows what it's like to be me". (Rogers, 1980: p.10)

Although I shall not give further consideration to the *Relational Defence* in this thesis, I shall, in chapter 3, defend the idea that empathy offers us a distinctive, and distinctively deep, way of understanding of other people in so far as it gives us knowledge of what their experiences are like. My priority will be to strengthen the Epistemic Defence by showing the important role that this knowledge can play in moral deliberation. A follow-up project could be to strengthen the Relational Defence by discussing how this knowledge can also be used to deepen our personal connections.

Before proceeding to explain in more detail how I intend to develop the Epistemic Defence of empathy, it will be helpful to first motivate my work by identifying two challenges that have been posed for the Epistemic Defence of empathy. I will then go on to show how my approach can help meet these challenges.

5. Two Challenges for the Epistemic Defence

5.1 What is special about empathy?

One challenge for the Epistemic Defence of empathy is to explain what is special about empathy. The critic of empathy can agree that attributive knowledge about the affective states of other people is helpful for moral deliberation, but point out that empathy is a flawed way of acquiring this knowledge. Empathy is flawed, the critic will say, because it is emotionally costly and because it can lead our decision-making astray. Accordingly, the critic will continue, if we want to acquire attributive knowledge, we should instead use testimony or inference.

In considering this challenge, it is crucial to remember that I am using the word 'empathy' to refer to a form of affective matching. One might also use that word in a broader way, such that it could refer to any means of acquiring attributive knowledge about the affective states of other people. On that broader usage of the term 'empathy', I could empathize with someone by listening attentively to their testimony, or by perceptively drawing inferences about what they are feeling, even if I do not in any way feel what they are feeling. On the narrow usage of the term 'empathy' that I am employing here, listening to a person's testimony or acquiring attributive knowledge about them through inference do not count as empathizing with them. On this approach, one can come to know a great deal about what another person is feeling without ever empathizing with them.

This challenge to the Epistemic Defence can be illustrated with an example. Suppose a friend is struggling after the death of a relative. In this case, having attributive knowledge of what the friend is feeling will be helpful in order to

know how to offer them sensitive support. Acquiring this attributive knowledge through empathizing with them, however, will involve feeling difficult (and potentially draining) feelings like grief, sadness, distress and fear. It may also lead you to think that you should make helping this friend your first priority, when in fact you may have more important concerns (such as caring for other friends, or for your family, or for your own basic needs). Instead of empathizing with this friend, the critic of empathy can say, you should find out about how they are feeling through an alternative means (e.g. by asking them how they are feeling).

To date, two different responses to this challenge have been offered. First, it has been argued that there are some situations in which empathy is the only way of acquiring attributive knowledge about the affective states of another person (Masto, 2014: pp. 84-85). To put the same point differently, it has been argued that there are some situations in which we cannot learn about the affective states of another person through testimony (e.g. if we cannot ask them, or if they cannot tell us) or through inference, and must therefore rely instead on empathy.

Secondly, it has been argued that empathy offers an epistemic output that is richer than that offered by testimony and inference. Song (2015: pp. 443-446) has suggested that whilst testimony and inference give rise to knowledge of what another person is feeling, only empathy can give us understanding of what they are feeling. For Song, understanding what another person is feeling involves knowing how what they are feeling relates to their other mental states, how it relates to their situation, and how it would be appropriate for us to respond to what they are feeling.

Although there is value in both of these responses, I think that there is a different, and more powerful, way of defending the idea that empathy has a distinctive epistemological contribution to make to moral deliberation. Whilst Song is correct in claiming that empathy offers a distinctively rich epistemic output, I do not think that this richness consists in the understanding that Song describes. Instead, I think that the distinctive importance of empathy lies in its capacity to give us phenomenal knowledge of what another person's affective state feels like. Before moving on to explain this idea in more detail, I now turn to consider the second challenge for the Epistemic Defence.

5.2. The Spotlight Effect Objection Revisited

Earlier, I summarized the Spotlight Effect Objection to empathy, which said that empathy for a suffering person can lead us to have excessive concern for that person, and to mistakenly prioritize helping them. With this objection in mind, we could ask anyone making the Epistemic Defence of empathy the following question: how can you reconcile your confidence in empathy as an epistemic tool with the evidence that it leads to errors in decision-making? As far as I can see, this is not a question that the existing proponents of the Epistemic Defence have engaged with.

To help clarify this discussion, it will be useful to work with an example. Consider, then, the following fictional case which is based on a case that was presented to subjects in one of Batson's psychology experiments (Batson, 1995). Prinz cites this case to support his development of (what I call) the Spotlight Effect Objection (Prinz, 2011a: p. 226).

The Sheri Case: Sheri Summers is a ten year old girl with a slow-progressing terminal illness. She is on the waiting list for medical treatment (which will presumably ease her pain, and perhaps even prolong her life). The other people on the waiting list are all children with terminal illnesses, and they have been ranked according to their priority for treatment. This ranking has taken into account how severe their illness is, and how long they have already been waiting for treatment. Sheri is not at the top of the list.

In considering this case, one question we can ask is: whom do we have most reason to help first? According to Prinz (and Batson), we have most reason to help the person at the top of the list. They are at the top of the list precisely because, all things considered, they are the person that we should help first. Accordingly, if anyone makes the moral judgement that the person that they have most reason to help is Sheri (who is not at the top of the list), then this person would be wrong.

Prinz's complaint about empathy is that it leads people to make such a judgement. In Batson's experiment, all of the participants are presented with a short vignette about Sheri, and are later given an opportunity to move Sheri up the priority list for treatment. Half of the participants are encouraged to 'stay objective' as they listen to Sheri's story, and of those participants, 33% of them chose to move Sheri up the priority list. The other half the participants are encouraged to empathize with Sheri (they are instructed to 'imagine Sheri's feelings') and of those participants, 73% chose to move Sheri up the list.

Let us call a judgement about what we have most reason to do a *global normative judgement*. I will soon contrast this with a *particular normative judgement*, which is a judgement about the strength of reason that we have to do a particular action. We can now say that, according to the Spotlight Effect Objection, empathy generates errors in global judgements. In other words, the objection says, empathy makes us think that we have the most reason to help the person that we have empathized with when there are, in fact, other people that we have greater reason to help.

It is important to remember that the Spotlight Effect Objection is not inconsistent with the basic claims of the Epistemic Defence of empathy. Those claims are: first, that attributive knowledge (about the affective states of others) is an important input into moral deliberation and, second, that empathy is a way of acquiring this attributive knowledge. In the Sheri case, the first claim implies that knowing about the affective states of the children on the waiting list (and how those states could be changed through medical treatment) is an important input into moral deliberation about which child to treat first. The extent to which a child is suffering is presumably one of the factors that helps us determine the extent to which we should prioritize them for treatment. The second basic claim of the Epistemic Defence says that empathy is one way of learning about the affective states of the children. For example, when we empathize with Sheri (perhaps through taking up her perspective, or through affective contagion), we can learn more about what she is feeling (perhaps: a mixture of physical pain, fear, hopelessness, isolation from friends, and so on). This information is useful in determining the strength of reason that we have to help Sheri.

According to the Spotlight Effect Objection, however, what really matters in cases such as these is making a correct global normative judgement about what we have most reason to do. The problem with empathy, according to that objection, is that it leads to mistaken global normative judgements. How might empathy do this? One possible explanation is that in empathizing with one person, the empathizer forgets the broader landscape of reasons that they face. For example, the empathizer, by focusing on the suffering of Sheri, may correctly believe that they have a strong reason to help Sheri, but forget that they have stronger reasons to help the children who are higher up on the priority list, who are also suffering. It is as if empathizing with Sheri makes the participants in Batson's experiment lose sight of these other children.

If this is the full explanation of how empathy leads us astray, then the defender of empathy can make the following response. They can say that the problem here is not intrinsic to empathy, but is rather in how empathy is integrated into moral deliberation. They can accept the Spotlight Effect Objection, and agree that one should not base one's decisions solely on the information gathered through a single usage of empathy, for that would lead one to neglect the broader range of considerations that should be taken into account. Instead, the defender of empathy can maintain that the empathizer should recognize the insight that they gain through empathy, and then take a step back to integrate that insight into a broader assessment of the possible actions they could take, and of the reasons for those actions. Thus, the empathizer in the Sheri case could say: 'yes, through empathizing with Sheri I now know how much she is suffering, and how important it is to help her. But I also know that the other children on the list are also suffering, and that some of those children are suffering even more than

Sheri, so I should begin by treating the child that has been deemed to be most in need'. In this way, an empathizer can avoid making glaring errors in their global normative judgements.

At this point, one might wonder what value empathy is adding here, if we already have priority list which tells us which child to treat first. I offer two responses to this. First, in an alternate version of the Sheri scenario, the priority list might be wrong. It could be that Sheri should be at the top of the list, when she is not. In that case, by empathizing with Sheri, and realizing the severity of her suffering, we could correct this error. Secondly, it could be that the overall treatment program is underfunded such that even if the priority list is correct, it will not lead to enough children being treated. In this case, empathizing with children on that list could lead us to realize that more resources should be allocated to treating those children.

In response to the Spotlight Effect Objection, then, the defender of empathy can acknowledge that many of the moral decisions that we face are complex, and that they involve large numbers of people and many competing considerations. They can maintain that one important consideration in our moral deliberations is the suffering of other people, and that empathy is a way of finding out about that suffering. To avoid errors in our global normative judgements, however, they can clarify that the empathizer must always remember that the suffering of the person that they are empathizing with is just one relevant consideration amongst many.

Although I think this is a promising reply to the Spotlight Effect Objection, it is currently inadequate. As it stands, it still does not explain why we should

empathize with other people instead of using testimony or inference to find out about their suffering. If those alternative means of acquiring attributive knowledge avoid the Spotlight Effect Objection, then, the critic of empathy can say, we should use those means instead of using empathy.

Furthermore, the critic of empathy could counter my above reply by saying that my earlier diagnosis of how empathy leads us astray was incomplete. The critic could say that, whilst it is true that the empathizer tends to lose sight of the big picture in moral deliberation, there is also a more fundamental problem with empathy in so far as it leads us to overstate the strength of reason that we have to help the person that we are empathizing with.

When we empathize with Sheri, and come to know about her suffering, we make the judgement that her suffering provides us with a reason (of some degree of strength) to help her. As I indicated above, I shall call this a *particular normative judgement*. A critic of empathy can accept that it is true that Sheri's suffering gives us a reason to help her. The critic can contend, however, that empathy leads us to make an exaggerated assessment of how strong this reason is. More generally, the critic can say that empathy leads us to make exaggerated particular normative judgements, in so far as the empathizer will overstate the strength of reason that they have to help the person that they are empathizing with. If this was so, then, even if the empathizer retains an accurate assessment of all of their other reasons for action, they could still come to a false global normative judgement about what they have most reason to do, because they have over-estimated the strength of reason that they have to help the person that they have empathized with.

In the Sheri case, the critic can say, empathizing with Sheri can lead you to make a false global normative judgement, even if you have correctly assessed your reasons to help the other children, because your empathy with Sheri will lead you to make an exaggerated particular normative judgement about the strength of reason that you have to help Sheri. We could even imagine a case where, prior to empathizing with Sheri, the empathizer already knew both what Sheri was feeling and the extent to which she was suffering. The critic of empathy could contend here that empathizing with Sheri would not provide us with any new information, but it would lead us to erroneously inflate our estimation of how strong our reason was to help Sheri.

If empathy systematically led us to make errors in our particular normative judgements, then there would be a strong case for avoiding empathy altogether. It would be better to use other ways of learning about the mental states of other people. The Spotlight Effect Objection then, as things stand, constitutes a major concern for the Epistemic Defence of empathy.

6. The Normative Epistemic Argument

In this thesis, I aim to strengthen the Epistemic Defence of empathy, in part by helping it to meet the two challenges outlined above. I shall do this by arguing that empathy is, in addition to being a source of attributive knowledge, a source of normative knowledge. In saying that empathy is a source of normative knowledge, I mean that it gives us knowledge of our reasons for action. To be more specific, I shall defend *The Normative Epistemic Claim*.

The Normative Epistemic Claim: When a person A empathizes with a person B who is feeling an affective state S, where S is an incidence of suffering, A will come to know that S gives A a reason to help B, and A will come know how strong that reason is.²⁹

It will be noted that this claim focuses on the case of empathizing with a person who is suffering. I shall restrict my focus in this thesis to such cases to make the discussion manageable, and because I think that such cases are important. Nevertheless, it is worth noting that I think that empathy can also play an important role in our deliberations about how to avoid inflicting harm on others (e.g. we could empathize with how a friend would feel if we failed to invite them to an event) and in our deliberations about how to promote joy in others (e.g. we could empathize with how a friend would feel if we invited them to an event). Indeed, at the end of this thesis I shall suggest that to recognize the full strength of reason that we have to help a suffering person, we should empathize not only with their present suffering (thus learning about our reasons to alleviate such suffering), but also with their potential joy (thus learning about our reasons to help them achieve such joy).

In defending the Normative Epistemic Claim, I demonstrate what is special about empathy as a way of learning about the affective states of other people. Whilst there are other ways, such as testimony and inference, to learn about such affective states, empathy is distinctive in so far as it also provides us with

²⁹ When I defend the Normative Epistemic Claim in chapter 5, I will give it a more precise formulation which employs the notion of a purely phenomenal reason. A purely phenomenal reason to alleviate another person's suffering is the reason that is given to us solely by the phenomenology of that person's suffering. I will say that empathy with a suffering person gives us knowledge of strength of the purely phenomenal reason that we have to help them.

knowledge about the reasons for action that are given to us by those affective states.

The Normative Epistemic Claim vindicates empathy as an input into moral deliberation. It directly rebuts the idea, considered above as part of the Spotlight Effect Objection, that empathy leads the empathizer to exaggerate the strength of the reason that they have to help the person they are empathizing with. According to the Normative Epistemic Claim, the empathizer will accurately assess the strength of the reason that they have to help the person that they are empathizing with. In other words, the empathizer will form a correct particular normative judgement.

To be clear, it is still possible that the empathizer might form an incorrect global judgement about what they have most reason to do, because they might fail to consider the reasons that they have for helping other people. To avoid this possibility, the empathizer must remember to step back from considering the person that they are empathizing with, so that they can integrate what they have learnt from empathy into a holistic process of moral deliberation.

If the Normative Epistemic Claim is correct, then it has important implications for ethics. I think that what the Spotlight Effect Objection gets right is that, generally, speaking, empathizing with a suffering person leads us to make an upward revision in our assessment of how strong our reason is to help that person. The Spotlight Effect Objection goes wrong, however, in suggesting that this is an error. According to the Normative Epistemic Claim, empathy leads us to the correct view about how strong our reasons are to help the person that we are empathizing with. Accordingly, if it leads us to an upward revision in that view, it

means that our reasons to help the other person are stronger than we previously thought. Indeed, in chapter 5, I defend the view that we have strong (stronger than we normally think) reasons to alleviate the suffering of other people. Empathy can have a profound impact on us precisely because it gives us knowledge of these reasons.

In practical terms, empathy is crucial in four ways. First, it can educate someone out of egoism. By egoism, I mean the view that the only reasons that we have are reasons to promote our own interests. On my approach, if an egoist empathized with a suffering person, they would learn that, contrary to the edicts of egoism, they had reasons to help that person by alleviating their suffering.

Second, empathy can educate someone out of prejudicial views on which they think that the suffering of certain groups of people is irrelevant. For example, a racist might discount the suffering of people within a particular racial group. If the racist could be brought to empathize with a person from that racial group, then the racist would learn not just that the person is suffering, but that their suffering matters. The racist will learn that their suffering matters in the sense that they, the racist, have a strong reason to alleviate it. Accordingly, my view may explain some of the value of empathy in conflict situations in which each party in the conflict tends to deny the significance of the suffering endured by the other parties in the conflict (Rosenberg, 2003).

Third, if a person believes that they have reason to help alleviate the suffering of others, but believes that these reasons are weak, then empathy will teach them that the reasons are strong. In this way, it is conceivable that a person

who empathizes extensively with the suffering of a group of people might be led to dedicate their life to helping that group.³⁰

Fourth, even if a person was generally committed to the view that they had strong reasons to alleviate the suffering of others, it may be possible for them to get into states of mind where they lose this commitment. In such a scenario, empathy can remind them of it. For example, I suspect that when I get into conflict with other people, I can know that they are suffering, but fail to acknowledge that their suffering gives me a reason to help them. Instead, I may be busy thinking about whether their suffering is my fault, or I may be strategizing about how I can prevent myself from being attacked or criticised. Taking time to empathize with the other person can remind me that, as much as I have a reason to alleviate my own suffering, I also have strong reasons to alleviate the suffering of the other person (or at least, to do my best to stop inflicting further suffering on them).

Aside from seeking to vindicate the role of empathy in everyday moral deliberations, I also hope to offer evidence that empathy is a topic of central philosophical importance. If the Normative Epistemic Claim is correct, empathy should be considered to be an important tool of moral epistemology. Moreover, as I construct my argument for the Normative Epistemic Claim, I will show that empathy has fascinating connections with debates in the philosophy of consciousness and the philosophy of hedonism. As far as I know, little has so far been done to connect empathy with either of these debates.

³⁰ Hoffman has a fascinating discussion of witnessing, which he defines to be an act of empathy that is so powerful that it leads the empathizer to commit to life-long action focused on eradicating the kind of suffering that they empathized with (Hoffman, 2011: pp. 238-245).

7. The Plan for this Thesis

Although my primary aim in this thesis is to strengthen the Epistemic Defence of empathy, I begin (in chapter 2) by discussing the claim that empathy is a source of altruistic motivation. My objective here will not be to use this claim to vindicate empathy. Rather, I shall argue that, as long as we employ the standard models for thinking about empathy, we lack an adequate explanation for how empathy gives rise to altruistic motivation. The implication will be that we should reframe how we think about empathy in order to find an adequate explanation for empathy-induced altruism. I shall suggest that a promising way forward lies in the idea that empathy is a source of phenomenal knowledge. If empathy is a source of phenomenal knowledge, then empathy with a suffering person will not only tell us which affective states they are feeling, it will also tell us what those affective states feel like.

In chapter 3, I develop this idea by connecting the notion of empathy with recent debates about phenomenal knowledge that have been taking place within the philosophy of consciousness. Drawing on Frank Jackson's case of Mary coming to see red for the first time (Jackson, 1982), I clarify the distinction between attributive knowledge and phenomenal knowledge, and argue that the only way that we can acquire phenomenal knowledge of the affective states of others is by empathizing with them. As part of this argument, I shall offer a novel definition of empathy. Although this definition of empathy will follow the contemporary norm in taking empathy to be a kind of affective matching, it will differ in important ways from the accounts of empathy that have been popular in the recent debate. In short, I shall argue, to fully understand the epistemic significance of empathy, we need to adjust our understanding of what empathy is.

In chapter 4, I begin to consider the moral significance of phenomenal knowledge. After reviewing the contemporary debate about the badness of pain, I argue that if we know what a person's suffering is like (which we can only do through empathizing with them), we will know how bad their suffering is for them. In chapter 5, I extend this idea, and argue that if we know what a person's suffering is like, we will know the strength of reason that their suffering gives us to help them. This is the Normative Epistemic Claim and is the main conclusion of this thesis. At the end of chapter 5, I return to consider how the Normative Epistemic Claim can be used to explain empathy-induced altruism.

In chapter 6, I conclude the thesis by summarizing the importance of empathy as an input to moral deliberation. At the end of the chapter, I move beyond cases of empathizing with suffering to consider the importance of empathizing with the joys of other people.

Chapter 2: Empathy and Altruism

1. Overview

One commonly-held view about empathy is that it is a source of altruistic motivation, such that when we empathize with a suffering person, there will be an increase in our altruistic motivation to help them. Let us call this *The Empathy-Altruism Hypothesis*. A second commonly-held view about empathy is that it is a form of affective matching, insofar as it involves feeling an affective state (an emotion, mood or physical sensation) that is similar to the one felt by the person that is being empathized with. The purpose of this chapter is to argue that combining these two views makes it difficult to answer an important explanatory question: how does empathizing with a suffering person (that is, matching their affective state) produce altruistic motivation to help them?

In the next section (section 2) I outline this question in more detail. Then (in sections 3 and 4) I show that four different suggestions for how to answer this question do not work. Finally (in section 5), I propose a novel explanation of empathy-induced altruism: I suggest that empathy with a suffering person motivates us to help them because it gives us knowledge of what that person's suffering is like. I call this the *Phenomenal Epistemic Explanation* of empathy-induced altruism, and the subsequent chapters of this thesis will help to develop this explanation. The task of this chapter is to motivate this later work, by showing that there is an important question for it to answer.³¹

³¹ Of course, I cannot show that my proposed solution is the only way of explaining empathy-induced altruism. I shall critique four existing explanations, but I cannot rule out the possibility that there might be some other explanation of empathy-induced altruism that I have not considered.

My next step in this chapter will be to present the explanatory question in more detail. To begin (in 2.1), I offer an example that will help to clarify the subsequent discussion. I then (in 2.2 and 2.3) give further introduction to the two 'commonly-held views' mentioned above: that empathy is a source of altruistic motivation, and that empathy is a form of affective matching. Finally (in 2.4), I show how combining these two views gives rise to an explanatory question that is difficult to answer.

2. The Explanatory Question

2.1 Steven and Oscar

Oscar has been depressed for months. Although he is still enrolled as a university student, he has stopped going to most of his classes, and has fallen behind with his work.

Steven is on the same course as Oscar, and is one of Oscar's friends. Through talking with Oscar at a recent social event, Steven knows that Oscar is feeling depressed. He also overheard Oscar saying that he would appreciate it if his friends would be more pro-active in arranging to meet up with him for meals, and in organizing joint study sessions.

However, although Steven does feel some motivation to do these things, he has not yet done any of them. He does not know Oscar that well, and Steven has a lot of his own work to do. Steven also finds it difficult to spend time with Oscar when Oscar is feeling depressed.

So far, Steven has also not empathized with Oscar. Although Steven has had his own run-in with depression, it was several years ago, and it is uncomfortable for him to remember it.

One day, Steven encounters Oscar in the university cafeteria, and is struck by how drained Oscar looks. Steven is prompted to empathize with Oscar (perhaps, in part by drawing on his own memories of feeling depressed).

To show how this case leads to an interesting question, it is necessary to first present two commonly-held views about empathy. The first is that empathy is a form of affective matching. The second is that empathy for a suffering person increases our altruistic motivation to help them. I now discuss each in turn.

2.2 Empathy as Affective Matching

As noted in the introduction to this thesis, contemporary philosophers have tended to draw a distinction between empathy and sympathy along the following lines. Sympathy is taken to be a reaction to the plight of others that involves feeling for them.

Definition of Sympathy: Person A is sympathizing with person B, just if A feels sad, concerned or distressed for B because A perceives or believes that B is in need.

Empathy, by contrast is taken to be a kind of feeling with another person. It is a kind of affective matching, such that it involves feeling an affective state that is similar to the one that is felt by the other person. More formally, we can say that there is an Affective Match Condition on empathy.

The Affective Match Condition: For it to be the case that person A is empathizing with person B with respect to B's feeling S (an affective state), it must be that A feels S*, where S* is similar to S.

For example, empathizing with an angry person will involve feeling anger and empathizing with a distressed person will involve feeling distress.

Applying this distinction to the example of Steven and Oscar, we can say that it would be one thing to sympathize with Oscar (which would involve feeling sad, concerned or distressed for him) and another thing to empathize with him (which would involve having empathic feelings of depression).³² How might Steven come to empathize with Oscar? One way would be for Steven to consciously imagine being in Oscar's situation (i.e. missing many of his classes, falling behind with his work, and spending a lot of time alone). As part of this process, Steven might recall how he felt when he was in a comparable situation. I refer to this route to empathy as 'perspective-taking'. It is also possible that Oscar might just 'pick up' Steven's emotion automatically after perceiving Steven's emotional expressions. I refer to this route to empathy as 'affective contagion'.³³ In the terminology that I am using here, perspective-taking and emotional contagion are different matching mechanisms through which an affective match can arise.

As I explained in the previous chapter, contemporary philosophers have disagreed about whether a particular matching mechanism is a necessary part of

³² For further explanation of the difference between empathy and sympathy, see section 2.1 in the previous chapter.

³³ For more detail on perspective-taking and affective contagion (including a brief review of the psychological literature that discusses how contagion works), see section 2.2 and section 2.3 in the previous chapter.

empathy. They have also offered different accounts of how to specify the similarity component of the Affective Match Condition (how similar do S and S* need to be?) and of whether there is also a knowledge condition for empathy (does the empathizer need to know what the other person's affective state is?). I shall discuss these issues further in the next chapter, where I develop my own account of empathy. For now, I wish to leave things open: all that I have said about empathy is that it is a kind of affective matching.

2.3 The Empathy-Altruism Hypothesis

A second claim that contemporary philosophers have often made about empathy is that it is a source of altruistic motivation.³⁴ I call this claim the *Empathy-Altruism Hypothesis*.

The Empathy-Altruism Hypothesis: If a person A empathizes with a suffering person (B), there will normally be an increase in A's altruistic motivation to help B.³⁵

This hypothesis has been supported with experimental evidence by social psychologist Daniel Batson (1991, 2011).³⁶ If applied to the case of Steven and

³⁴ See Darwall (1998: pp. 273-274), Nussbaum (2001: pp. 327-341), Maibom (2006: pp. 162-164), Slote (2011: p. 20), Simmons (2014), and Denham (2017: p. 233).

³⁵ Although, within both the philosophical and psychological literature cited here, the examples used to illustrate the Empathy-Altruism Hypothesis always involve empathizing with a suffering person, the standard way of formulating that hypothesis is as saying that altruism will be generated whenever we empathize with a person who is in need. The reason that I do not use that formulation is that a person may be in need without suffering (e.g. if that person is unaware of their need, or is in denial of it). We may come to be altruistically motivated to help such a person, but I am sceptical that empathizing with them would give rise to such motivation.

³⁶ Batson's work extends (what he takes to be) the prevailing view in psychology that empathy generates pro-social motivation (see, e.g., Eisenberg and Miller, 1987). Although contemporary philosophers have generally been impressed by Batson's work (see the references above), it has been subject to criticisms from other psychologists, who contend that although Batson has provided further evidence that empathy leads to an increase in helping motivation, he has not

Oscar, this hypothesis implies that when Steven empathizes with Oscar, there will be an increase in Steven's altruistic motivation to help Oscar.³⁷

What is meant by altruistic motivation here? There are several ways of defining altruism (for a review, see Batson, 2011: pp. 23-29), and here I shall adopt Batson's definition.

Definition of Altruistic Motivation: Altruism is 'a motivational state with the ultimate goal of increasing another's welfare.' (ibid.: p. 20)

Batson contrasts altruistic motivation with egoistic motivation, which seeks to increase the acting agent's own welfare as its ultimate goal. If Steven wanted to help Oscar to impress his other friends, or to feel good about himself, then he would be motivated egoistically. In saying that Steven's motivation to help Oscar is altruistic, I mean that he is acting for the sake of Oscar.³⁸

In saying that empathy 'normally' gives rise to altruistic motivation, I mean that it will do so in the vast majority of cases but I allow that there may be exceptions.³⁹ Later in this chapter, I shall discuss one possible exception case: if the empathizer gets over-aroused by empathizing with a suffering person, they can become focused on alleviating their own distress, rather than on helping the person that they are empathizing with. Another possibility is that a person might calmly empathize with another person's suffering, but still experience no increase

done enough to demonstrate that this helping motivation is altruistic. See, for example, Hoffman (1991) and Cialdini (1997). For a response, see Batson (2011: ch. 6).

³⁷ I am assuming here that the case of Steven and Oscar is a 'normal' case. I say more on what I mean by this below.

³⁸ A person's action might be driven by a mixture of altruistic and egoistic motivations. The Empathy-Altruism Hypothesis predicts that empathy will increase a person's altruistic motivation, but it does not say that altruistic motivation will be the only kind of motivation that they have (they may still have egoistic motivations for helping the other person).

³⁹ For a stronger view, which does not allow for exceptions, see Simmons (2014). I think that Simmons' view is mistaken, because it relies on the flawed Intrinsic Motivation explanation for empathy-induced altruism. I critique that explanation in section 3.3 below.

in altruistic motivation to help them.⁴⁰ In general, by exploring how empathy gives rise to altruistic motivation, I aspire to be in a better position to specify more precisely the conditions under which empathy does give rise to altruistic motivation.

Just as the Empathy-Altruism Hypothesis does not say that empathy is sufficient for altruistic motivation, it also does not say that it is necessary for such motivation. It is compatible with the possibility that altruistic motivation can arise without the presence of empathy.⁴¹

Caution is required in citing Batson's research in support of the Empathy-Altruism Hypothesis. As I noted above, contemporary philosophers have typically defined empathy to be a form of affective matching, but Batson defines empathy differently. Batson defines empathy to be what I am calling sympathy.

Batson's Definition of Empathy: Empathy is 'other-oriented emotion elicited by and congruent with the perceived welfare of someone in need' (Batson, 2011: p. 11).⁴²

Such other-oriented emotions might include: 'feelings of sympathy, compassion, softheartedness, tenderness, sorrow, sadness, upset, distress, concern, and grief.' (ibid.). In saying that these emotions are other-oriented, Batson means that they are felt for another person, as opposed to being felt for oneself. The requirement that the emotions are congruent with the perceived welfare of

⁴⁰ For an interesting discussion of this scenario, see Nussbaum (2001: p. 228-230).

⁴¹ For an account of altruistic motivation that does not reference empathy, see Nichols (2001).

⁴² It should be noted that Batson uses the term 'empathy' interchangeably with the term 'empathic concern'.

someone in need simply means that if we perceive that the person is in need, we must feel a negative emotion, such as one of the emotions listed above.⁴³

Thus, in the terminology that I am using here, what Batson calls ‘The Empathy-Altruism Hypothesis’ is not a claim about the empathy-altruism relationship, but rather a claim about the sympathy-altruism relationship. Batson’s conclusion could be relabelled as the *Sympathy-Altruism Hypothesis*.

The Sympathy-Altruism Hypothesis: If a person A sympathizes with a person in need (B), there will be an increase in A’s altruistic motivation to help B.

Nevertheless, several of the philosophers cited above have pointed out that while Batson’s primary focus is on the sympathy-altruism relationship, his research also provides evidence for the Empathy-Altruism Hypothesis. I shall discuss this issue in section 4, when I return to consider the relationship between empathy and sympathy in more detail.

2.4 The Explanatory Question

If the Empathy-Altruism Hypothesis is correct, a question arises about how empathy leads to an increase in altruistic motivation. I call this the *Explanatory Question*.

The Explanatory Question: how does empathizing with a suffering person lead to an increase in altruistic motivation to help that person?

⁴³ Batson would allow that we can also empathize with people whom we perceive to be flourishing. In these cases, a congruent emotion would be a positive one, such as joy or gratitude. Batson does not claim that empathy in such cases would produce an increase in altruistic motivation.

Applying this question to the case above, we can ask: how does Steven's empathy for Oscar lead to an increase in Steven's altruistic motivation to help Oscar? In the next two sections (3 and 4), I argue that this question has not yet been satisfactorily answered.

It should be noted that this question is only difficult to answer when we define empathy to be a form of affective matching. If, by contrast, we define empathy to be sympathy (as Batson does), then the question is much easier to answer. One can simply say (as Batson does) that sympathy is a type of emotion, and that one of the properties that this type of emotion has is that it is altruistically motivating.⁴⁴

As we shall see, a similar move cannot be made to explain how empathy (as affective matching) increases altruistic motivation. In empathy, the empathizer can feel a wide range of affective states, depending on what the other person is feeling. These affective states can have a wide range of motivational properties, and many have no obvious link to altruistic motivation. Consider, for example, empathizing with a person's headache, or with their guilt, or, as in the case of Oscar, with their depression. When Steven empathizes with Oscar, and thereby comes to be altruistically motivated to help him, there is thus a puzzle: how does feeling depression increase Steven's altruistic motivation to help Oscar? Is depression not the kind of thing that is demotivating?

⁴⁴ In general, Batson think emotions can play an 'amplificatory function', which means that they can increase the intensity of our goal-directed motivations. Batson proposes that sympathy (what he calls empathy) has such a function. What makes sympathy a special case is that the goal it motivates us towards achieving is promoting the welfare of another person (rather than promoting our own welfare). For Batson, the claim that sympathy is altruistically motivating is not an analytic one (it is not supposed to be part of the definition of sympathy). Rather, it is an empirical claim that Batson supports with evidence. For further discussion, see Batson (2011, p. 31).

I now consider four different strategies that have been offered for answering the Explanatory Question, and argue that each of them is inadequate. They are inadequate because they cannot explain empathy-induced altruism in a full range of cases. In other words, for each strategy, there are cases of empathy-induced altruism that the strategy cannot explain.⁴⁵

In section 3, I deal with the first three of these strategies: the *Standard Epistemic Explanation*, the *Aversive Arousal Reduction Explanation*, and the *Intrinsic Motivation Explanation*. Section 4 considers, and critiques, the fourth and final strategy: the *Empathy-Sympathy Explanation*. I give this strategy its own section both because it requires a more extensive discussion, and because it integrates two of the preceding strategies.

3. Three Attempts to Answer the Explanatory Question

3.1 The Standard Epistemic Explanation

One possible way to explain empathy-induced altruism is (what I call) the *Standard Epistemic Explanation*. This explanation says that empathy for a suffering person gives rise to altruistic motivation to help them because it teaches us about what that person is feeling. On this approach, altruistic motivation will arise provided two conditions are met. The first condition is that empathy for the suffering person yields new knowledge about which affective states that person is feeling. The second condition is that the empathizer cares about the other person, where such care involves having a disposition to be altruistically motivated to help

⁴⁵ One might reply that even if the strategies are individually inadequate, they might collectively explain the full range of cases of empathy-induced altruism. As I explain further below, the example of Steven and Oscar shows that this is not the case.

that person when presented with the information that they are suffering. If both of these conditions are met, then when I empathize with a suffering person, I will learn about which affective states they are feeling and, because I care about them, I will thus come to be altruistically motivated to help them. For example, suppose that one of my friends (whom I care about) has lost his job. Until I empathize with him, I might not realize how disappointed and fearful he is. When I empathize with him, and thus come to learn how he is feeling, I will thereby come to be altruistically motivated to offer him support.

Why should we think that empathy with a suffering person will give us knowledge of what they are feeling? One possibility is that such knowledge is a necessary part of empathy, such that one cannot be said to be empathizing with another person unless one knows what they are feeling. Another possibility is that such knowledge (though not a necessary feature of empathy) often arises when we match the affective states of other people. The explanation for this would be as follows: (i) empathy involves matching the affective state of another person, (ii) if you have matched the affective state of another person, and you know both that you have done so and what your resultant affective state is, then (iii) you will be in a position to infer what the other person is feeling.⁴⁶ For example, if I am empathizing with a friend John who is feeling sad, I will (by the Affective Match Condition) be feeling sad. If I know that I am feeling sad, and if I know that my

⁴⁶ It should be noted that on some accounts of empathy (see, e.g., Sober and Wilson, 1998: pp. 234-235), the empathizer first knows what the other person is feeling and then, through having this knowledge, comes to feel as the other person feels. This is also a plausible interpretation of Hume's account of how empathy (which he refers to as 'sympathy') comes about (see my discussion of Hume in section 2.3 of the previous chapter). On these accounts, the shared affective state is redundant in so far as it is not part of the causal process that gives rise to altruistic motivation.

sadness has arisen through empathizing with John, then I will be in a position to infer that John is feeling sad.

This approach, however, leaves room for the possibility that the empathizer fails to know what the other person is feeling: the empathizer may fail to identify what they themselves are feeling, or they may fail to be aware that they have that feeling because they are empathizing with someone else. These worries may apply in particular to cases of empathy arising through affective contagion. It is also important to note that a would-be empathizer may also go wrong in believing that they have successfully matched the affective state of another person when, in fact, they have not.⁴⁷ However, this is not a problem with genuine empathy (which requires a successful match of affective states), but rather with our ability to achieve genuine empathy, and to be able to distinguish between cases where we have achieved such empathy and cases where we have not.

Before moving on, it is worth pointing out that the idea that empathy gives us knowledge of what others are feeling is closely related to the view within the mindreading debate that simulation is a crucial mechanism through which we come to know the mental states of others (Gordon, 1996; Ravenscroft, 1998; Goldman, 2006). Roughly speaking, the term 'simulation' has here been used to refer to the process of imagining being in another person's situation, and observing which psychological states one thereby comes to be in. What is in that debate called 'simulation' thus bears a close connection to what I am calling 'perspective-taking'. Given this close connection, I wish to suggest that the

⁴⁷ There is a large scientific literature exploring biases involved in perspective-taking (one of the mechanisms that can give rise to an affective match). For example, see Van Boven & Loewenstein (2003). For an overview, see Goldman (2006: ch.7).

epistemic defences of simulation lend further support to the idea that empathy (or at least, empathy which arises through perspective-taking) can give us knowledge of what other people are feeling.

According to the Standard Epistemic Explanation, the second condition that must be fulfilled in order for empathy to give rise to altruistic motivation is that the empathizer must care about the person that they are empathizing with. If, by contrast, the empathizer is utterly indifferent towards the other person, then finding out about the suffering of that person will not motivate the empathizer to help them. In such a case, empathy would fail to give rise to altruistic motivation. Where the empathizer does care about the person that they are empathizing with, the amount of altruistic motivation generated will be proportionate to the strength of care that the empathizer has for the other person. Empathizing with a close friend or family member would generate more altruistic motivation than empathizing with a stranger.

It is worth pointing out that, if the Standard Epistemic Explanation gives the full story about empathy-induced altruism, it may be misleading to say that empathy generates altruism. Although, on this approach, empathy can cause an increase in altruistic motivation, it can do so only by interacting with a prior altruistic disposition, which is the care that the empathizer has for the person that they are empathizing with. This disposition is not itself caused by empathy.

Although the Standard Epistemic Explanation might offer part of the story of how empathy gives rise to altruistic motivation, it faces an important problem. To see this, note that the explanation does not work in the case of Steven and Oscar. This is because, prior to empathizing with Oscar, Steven already knew

what Oscar was feeling (Oscar had told Steven that he was depressed). If the knowledge of Oscar's emotional state was going to impact Steven's motivations, it would have done so already. The implication is that if empathy increases Steven's motivation to help Oscar, this increase in motivation cannot be explained by the standard epistemic function of empathy.⁴⁸ More generally, although the standard epistemic function might explain some of the impact that empathy has on our motivation to help others, it cannot explain all of it. It does not work in cases where the empathizer already knew what the other person was feeling.

3.2 The Aversive Arousal Reduction Explanation

How else might empathy motivate us? I now consider two ways in which empathic affect can directly motivate the empathizer to help the person that they are empathizing with. The first is the *Aversive Arousal Reduction Explanation*.⁴⁹ It posits that (i) when we empathize with a suffering person, our empathic affective state will be unpleasant, (ii) when we have an unpleasant affective state, we will be motivated to eradicate it, and (iii) one of the ways of eradicating an unpleasant empathic affective state is by helping the person that we are empathizing with.

Much of this seems plausible to me. When I experience unpleasant affective states (such as a headache, stress, or loneliness), I am typically

⁴⁸ One might try to defend the Standard Epistemic Explanation by maintaining that in the case of Steven and Oscar, Steven's empathy for Oscar would not increase Steven's altruistic motivation to help Oscar. More generally, one might maintain that in cases where the empathizer already knows what the other person is feeling, empathy will not lead to an increase in altruistic motivation. My view is that empathy does lead to an increase in altruistic motivation in these cases, and that there is consequently an important question about how we are to explain this increase. To some extent, my view is based on my reflections on my own experiences of empathy. My view is also supported by the Normative Epistemic Claim that I defend in chapter 5.

⁴⁹ This terminology is borrowed from Batson's discussion of how sympathy (which he calls empathy) might motivate us (Batson 2011: ch. 3), although I have not seen anyone explicitly claim that this is how empathy motivates us.

motivated to try to eradicate them. I might do this by taking a pain relief drug, by distracting myself (e.g. by watching TV), or by trying to make an improvement in my external circumstances (e.g. by spending more time with friends). When I empathize with someone else who is feeling an unpleasant affective state, I will feel a similar (and similarly unpleasant) affective state. My empathic affect will be unpleasant and, according to the above intuition, I will be motivated to try to eradicate it. One way I could do this is by helping the person that I am empathizing with: If the other person no longer feels an unpleasant affective state, then I will no longer have to suffer by empathizing with them.⁵⁰

Consider how this approach applies to the example of Steven and Oscar. In that example, Oscar feels depressed and this is unpleasant. When Steven empathizes with Oscar, he will also have feelings of depression, and these will also be unpleasant. Steven will thus be motivated to eradicate his unpleasant empathic feelings, and one way to do this will be to help Oscar feel better, so that empathizing with him no longer induces unpleasant feelings in Steven. Thus: Steven's empathy for Oscar will motivate Steven to help Oscar.

While this may explain one type of empathy-induced motivation, it cannot be an explanation of altruistic motivation, because the motivation it explains is explicitly egoistic. Steven is only helping Oscar to get rid of his own unpleasant feelings. Moreover, Steven could also eradicate his empathic affective state by finding a way to stop himself from empathizing with Oscar. For instance, Steven could end his unpleasant empathic feelings by walking away from Oscar, and avoiding him in the future. In short, although the Aversive Arousal Reduction

⁵⁰ I would only help the other person if I understood that my unpleasant feelings had arisen through empathizing with them.

Explanation might sometimes account for empathy-induced helping, it can never account for empathy-induced altruistic motivation.

3.3 The Intrinsic Motivation Explanation

Empathic affective states might motivate us in a different way. To see how, consider first how ordinary emotions motivate us. As I discussed above, ordinary emotional states might sometimes motivate us by being unpleasant, and thus by provoking our desire to eradicate unpleasant affective states. But this is not the only way that ordinary emotions motivate us. Sometimes, emotions have goals (or desires) built in to them. For example, if I am afraid of a snake, I will be motivated to secure my safety by getting more distance between myself and the snake (either by moving it, or by moving myself). Achieving this goal might also eliminate my fear, but eliminating the fear is not the goal. If someone told me that I could eliminate my fear by taking a sedative, this would not make me any safer. Where emotions have motivational goals built into them, I will say that they are intrinsically motivating.

To take another example, suppose that I am angry at someone for the way that they treated one of my friends, and that this anger is motivating me to want to punish that person. Achieving this goal (punishing the perpetrator) might dissipate the anger, but dissipating the anger is not the goal.

The idea that emotions can be intrinsically motivating is central in the work of Daniel Batson (although he does not use the terminology of 'intrinsic motivation'). As discussed above, Batson claims that sympathy (what he calls empathy) is intrinsically motivating. He needs to claim this because he wants to

say that sympathy motivates us altruistically. Accordingly, the motivating power of sympathy cannot be explained by the Aversive Arousal Reduction Hypothesis. That is, when we sympathise with another person, and are thereby motivated to help them, Batson maintains that this motivation does not arise from our (self-focused) desire to eradicate our unpleasant feeling of sympathy. For Batson, sympathy is intrinsically motivating, and it moves us to help others for their sake.

If the sympathy-altruism relationship can be explained by appealing to the idea of intrinsic motivation, can the same be said for the empathy-altruism relationship? Maibom's account of empathy-induced altruism might be used to suggest that it can (Maibom, 2007). The idea runs as follows. Suppose that in the usual case, our (non-empathic, non-sympathetic) emotions are intrinsically motivating in an egoistic way. That is, they motivate us to pursue goals for our own sake. Sadness might move us to get consolation from a friend, fear might move us to avoid a dangerous situation, anger might move us to defend ourselves against an aggressor, and so on. In each of these cases, we seek the goal for our own sake.

Now consider what happens if I empathize with another person who is feeling such an emotion. Because empathy requires that I match their affective state, I will necessarily feel an emotion that is similar to theirs. One way of interpreting this is that I must feel an emotion that has the same motivational properties as theirs.⁵¹ So: if their emotion motivates them to pursue a goal for

⁵¹ This appears to be Maibom's view. It relies on her definition of empathy, on which it is a requirement that the empathic emotion is felt for the other person (Maibom, 2007: p. 168). Thus, on her account, to empathize with a distressed friend, it is not enough for me to simply feel distressed; my distress must be felt for them (as opposed to being felt for myself). Given that the other person also feels distressed for themselves, it follows that my empathic distress is, in this regard, the same as their distress.

their sake, my empathic emotion will also motivate me to pursue that goal for their sake. Whereas their emotion motivates them egoistically, my empathic emotion motivates me altruistically. For example, if a friend feels sad, and is thus motivated to find consolation for their sake, then when I empathize with them, I will feel sad for them and will thus be motivated to console them for their sake.

Again, I think there is something in this line of thought. It cannot, however, give a general explanation of empathy-induced altruism. The reason for this is that it cannot account for empathy-induced altruism in cases where the affective state in question is either motivationally inert, or is motivationally destructive.

Return to the case of Steven and Oscar. When Steven empathizes with Oscar, the affective state that he will feel is depression. Perhaps he will also come to feel related affective states that Oscar might be feeling, like exhaustion, self-criticism and loneliness. These are states that are debilitating. If Oscar's depression is demotivating for Oscar, Steven's empathic feelings of depression will be demotivating for Steven (even if they are in some sense felt for Oscar). Indeed, they may even be destructive. If Oscar is feeling critical towards himself (perhaps we could call this guilt, or self-loathing), then Steven will, through his empathy for Oscar, come to feel critical towards Oscar. Yet this looks like the kind of feeling that might motivate Steven to harm Oscar, to deprive him of benefits, or to avoid him. It certainly does not look like the kind of feeling that will be altruistically motivating.

The Steven and Oscar case might be used by a critic of empathy to argue that empathy is a bad idea.⁵² The critic could say that in this kind of case,

⁵² See, for example, Prinz's scepticism about the motivational efficacy of empathy (Prinz, 2011a: p. 220).

empathic affect is debilitating, and detracts from altruistic motivation. One might think that it would be better if Steven stuck with knowing what Oscar was feeling, and did not get bogged down in sharing Oscar's feelings. Such sharing, one might think, will not increase Steven's motivation to help Oscar.

Reflections on my own experiences of empathy suggest the opposite. My view is that Steven's empathy for Oscar will increase his altruistic motivation to help him.⁵³ If such an increase occurs, then, I have argued, the Intrinsic Motivation Explanation cannot account for it (because Oscar's feelings are not constructively motivating). Indeed, neither can the Aversive Arousal Reduction Explanation (because it can only explain egoistic motivation), and nor can the Standard Epistemic Explanation (because Steven already knew what Oscar was feeling). A different explanation is required.

4. The Empathy-Sympathy Explanation

4.1 Overview

The final explanation that I shall consider focuses on the relationship between empathy and sympathy. I call it the *Empathy-Sympathy Explanation* and it has two parts. First, it claims that empathy is sufficient for sympathy.

⁵³ More generally, my claim is that empathizing with people who are experiencing debilitating emotional states can increase altruistic motivation to help them. Batson's experiments lend support to this view. Batson has consistently shown that subjects who take up the perspective of a suffering person (thereby empathizing with them) are more likely to help that person than subjects who are instructed to 'stay objective'. Crucially, these experiments are often set up in such a way that the suffering person is likely to feel a debilitating emotion, such as fear, depression or exhaustion. The implication is that empathy with respect to these emotions leads to an increase in altruistic motivation. For an example of one of Batson's experiments, see the discussion of 'Carol's Case' below (Toi and Batson, 1982). For a full summary of the scenarios used in Batson's experiments, see the appendices of Batson (2011).

The Empathy-Sympathy Hypothesis: When A empathizes with a suffering person (B), there will be an increase in A's sympathy for B.⁵⁴

Second, it claims that sympathy is altruistically motivating (above, I call this the Sympathy-Altruism Hypothesis).

Combining these two views yields an explanation of how empathy for a suffering person gives rise to altruistic motivation to help that person: the empathy leads to an increase in sympathy, which is altruistically motivating. Applied to the example of Steven and Oscar, this explanation says that when Steven empathizes with Oscar, Steven's sympathy towards Oscar will increase, and the additional sympathy will motivate Steven (altruistically) to help Oscar.

As we shall see, there is persuasive evidence for the Empathy-Sympathy Hypothesis. Earlier, I said the same for the Sympathy-Altruism Hypothesis. Thus, the Empathy-Sympathy explanation of empathy-induced altruism has a strong empirical backing. The problem with it, however, is that it does not explain enough. More specifically, the problem is that it does not explain how empathy gives rise to sympathy. In place of a question about how empathy gives rise to altruistic motivation, we are instead left with a question about how empathy gives rise to sympathy. In other words: we have a new Explanatory Question.

The Explanatory Question:* How does empathy for a suffering person lead to an increase in sympathy felt towards that person?

⁵⁴ This hypothesis does not say that empathy is necessary for sympathy, because it allows that sympathy can arise without empathy. Also, as I discuss below, the claim that empathy is sufficient for sympathy may be too strong, as there may be circumstances in which empathy arises without there being a corresponding increase in sympathy.

I suspect that the force of this question has been obscured by the view that empathic affect simply is sympathy. I call this view the *Constitutive View* of the empathy-sympathy relationship, and I shall argue that it is not plausible. Instead, we should adopt the *Causal View* of the empathy-sympathy relationship, on which empathy does not involve sympathy (as the empathic affect), but instead causes additional sympathy to arise. On the Causal View, however, we are left with an important question about how this happens.

To begin this section, I summarize the evidence for the Empathy-Sympathy Hypothesis. I then present, and critique, the Constitutive View, before showing how the alternative view (the Causal View) leaves us with an explanatory puzzle.

4.2 The Empathy-Sympathy Hypothesis

An important source of evidence for the Empathy-Sympathy Hypothesis is the experimental research of Daniel Batson (Batson, 1991 and 2011). As noted above, care is required in interpreting Batson's findings, because his terminology is different to that which tends to be employed in the contemporary philosophical literature. Batson uses the term empathy to refer to what most philosophers call sympathy, and thus it is helpful to relabel his primary conclusion (which he calls the Empathy-Altruism Hypothesis) as the Sympathy-Altruism Hypothesis.

To provide evidence for the Sympathy-Altruism Hypothesis, Batson has run dozens of experiments in which subjects are divided into two groups: a high sympathy group (which Batson calls a high empathy group) and a low sympathy group. Subjects are then presented with information about a stranger in need, before being given an opportunity to help that stranger. By splitting the subjects

into the high and low sympathy groupings, Batson can study the impact that sympathy has on a subject's motivation to help the stranger.

There are a variety of ways in which Batson creates this division, but one of them is to give the subjects different instructions for how to listen when presented with the stranger's problem. The low sympathy group is told to 'stay objective', whilst the high sympathy group is told to 'imagine how the other person feels'. According to both my terminology and Batson's terminology, this latter instruction induces subjects to engage in perspective-taking with the person in need.⁵⁵

To illustrate, let us consider one of Batson's experiments in more detail. In one study, participants (undergraduates) are presented with an audio recording about Carol, a fellow student who has been in a car accident, and who now has trouble getting into school and keeping up with her studies (Toi & Batson, 1982). At the end of the recording, participants have an opportunity to offer to help Carol with her work. As noted above, half the participants are asked to 'stay objective' as they listen to Carol, whilst the other half are asked to imagine how Carol feels (to take up her perspective).

In identifying the latter group as the 'high sympathy' group, Batson is relying on it being the case that perspective-taking produces sympathy. His results support this idea: those who are asked to take up Carol's perspective are much more likely to help her than those who are asked to stay objective. When it is easy for participants to avoid helping Carol (the 'easy escape condition'), 71%

⁵⁵ Other methods include: raising sympathy levels by making the subject perceive themselves to be more similar to the person in need (following Krebs, 1975), or lowering sympathy levels by giving subjects a drug (actually a placebo) which makes them interpret their feelings of sympathy as personal distress (Batson et al, 1981).

of the 'perspective-taking' group helped her, whereas only 33% of the 'stay objective' group did so. Batson's explanation of this is that the former group experienced higher levels of sympathy. Batson is here drawing on a body of social psychological evidence for the potency of perspective-taking which began with the work of Stotland (Stotland, 1969).

There is, thus, evidence for the idea that perspective-taking leads to an increase in sympathy. Can this idea be used to support the claim that empathy leads to an increase in sympathy? Following at least two contemporary philosophers (Nussbaum, 2001; Darwall, 1998), I suggest that it can, with the following rationale: perspective-taking is a mechanism that gives rise to empathy, and it is through giving rise to empathy that perspective-taking leads to an increase in sympathy. If this is right, then if empathy arises without perspective-taking (e.g. through affective contagion), it will also lead to an increase in sympathy.

So far, I have summarized how one might interpret Batson's work as offering evidence in favour of the Empathy-Sympathy Hypothesis. It is worth also noting that this hypothesis has a wider acceptance in the psychological literature about empathy. For example, consider the following excerpt from Martin Hoffman, a psychologist well-regarded for his research into the developmental origins of empathy:

From then on, through the last three empathy development stages and throughout adult life, empathic distress has a sympathetic component. This is crucial because the sympathetic component gives empathic distress a clearly pro-social dimension. (Hoffman, 2011: p. 234)

Indeed, Hoffman's point is not just that empathy gives rise to sympathy, but that the sympathy is somehow part of the empathy. I shall discuss this rendering of the Empathy-Sympathy Hypothesis further below.

Although the Empathy-Sympathy Hypothesis is widely accepted, there is an important clarification that is required. So far, I have presented it as if empathy always implies the presence of sympathy. This is misleading. A common qualification is that empathy can sometimes give rise to sympathy, but that in other situations it can manifest as personal distress (Hoffman, 1978; Spinrad & Eisenberg, 2013).⁵⁶ Personal distress is an emotional reaction that leads the subject to focus on helping themselves by soothing their own emotional state. They may still help the other person, but they will do so only as a way of alleviating their own distress.⁵⁷

The personal distress reaction is exemplified by very young children: although they can empathize with others, this empathy leads them to seek help for themselves, and not to direct concerned attention towards the people whom they empathize with.⁵⁸ In adults, empathy with others can still lead to personal distress when the empathic affect is very strong.⁵⁹ This might be called 'empathic over-arousal', and can lead to vicarious trauma (Kaplan, 2011). However, in cases of adult empathy that do not involve extreme emotional intensity, the

⁵⁶ One way to eliminate this possibility is to define empathy in such a way as to exclude the cases of affective matching that give rise to personal distress. For example, Maibom requires that the empathic affect be felt for the other person (Maibom, 2007). Personal distress is felt for oneself, and thus on Maibom's definition, the empathic affect can never be personal distress.

⁵⁷ See section 3.2 above, which discusses the Aversive Arousal Reduction explanation of empathy-induced altruism.

⁵⁸ For a succinct summary of this developmental stage, see Hoffman (2011: pp. 234-235).

⁵⁹ The threshold at which an empathic affect becomes 'too strong' (thus leading to personal distress) appears to vary from person to person, depending on their ability to regulate their emotions. For an interesting discussion of this point, see Spinrad & Eisenberg (2013).

Empathy-Sympathy Hypothesis can, as a general rule, be expected to hold. Batson's cases fall into this category.

If the Empathy-Sympathy Hypothesis is correct, and if sympathy is altruistically motivating (as the Sympathy-Altruism Hypothesis says it is) then we appear to have an explanation of empathy-induced altruism. The problem is that this explanation does not take us far enough. It does not answer the question of how empathy gives rise to increased sympathy.

To show that this question cannot be easily answered, I now distinguish two interpretations of the Empathy-Sympathy Hypothesis. On the *Constitutive View*, empathy gives rise to increased sympathy because empathic affect simply is sympathy. If this view is correct, then there is no puzzle about the Empathy-Sympathy Hypothesis. I argue, however, that it is not correct. Instead, we should favour the alternative *Causal View*, on which empathy does not itself involve sympathy, but causes an increase in sympathy. On that view, a puzzle remains about how empathy does this.

4.3 The Constitutive View

The first way of understanding the relationship between empathy and sympathy is the *Constitutive View*.

The Constitutive View: When a person A empathizes with a suffering person (B), the empathic affect felt by A is sympathy for B.

As I mentioned above, Hoffman appears to endorse this view. It is also one possible interpretation of Darwall's view of the empathy-sympathy connection,

although Darwall would distinguish between empathy and ‘proto-sympathetic empathy’ and would say that it is only the latter that involves sympathy (Darwall, 1998: pp. 270-272).⁶⁰ If this view is correct, then the Empathy-Sympathy Hypothesis follows as a matter of definition, and no further explanation of that hypothesis is required.

How might the Constitutive View work? Consider an example. Suppose a person feels sad, concerned or distressed for themselves. By the definition of sympathy given above, we can then say that they feel sympathy towards themselves. If another person empathizes with this person, then (by the definition of empathy as affective matching) they will come to feel an empathic affect that is similar to the other person’s affect. Thus: they will feel sad, concerned or distressed. If we place a stipulation on empathy that the match involved must be not just of the type of affect, but also of whom the affect is felt for, then it follows that the empathizer will feel sad, concerned, or distressed for the other person. In other words, the empathizer will feel sympathy for the other person, and this sympathy is the empathic affect. We could still say that empathy produces sympathy, but only in the sense that the matching mechanism (which produces the empathic affect) produces sympathy.

This view faces an obvious objection: it cannot be true in cases where the suffering person does not feel emotions like sadness, concern or distress, but

⁶⁰ Darwall’s view is complex and it is hard to summarize fully here. One possible reading is as follows. In empathy, we use emotional contagion or projection (what I call perspective-taking) to come to feel as the other person is feeling. In the case of proto-sympathetic empathy, we empathize, and we also put our attention on the phenomenology of the other person’s experience. This will guide our empathic feelings towards being felt for the other person (we might think ‘it is awful that they are suffering in this way!’), rather than being felt for ourselves. In short, empathy with a distressed person will make us feel distress and if we also attend to what the distress is like for the other person, this will ensure that our distress is felt for the other person.

instead feels emotions that have no obvious connection with sympathy.⁶¹ In the example above, Oscar felt depressed, and perhaps also lonely, helpless and self-critical. These are not emotions one would normally associate with feeling sympathy for oneself. Through his empathy with Oscar, Steven will come to feel these emotions. Steven's empathic affective states will not include sympathy towards Oscar and therefore this case shows the Constitutive View to be false. It is also hard to see how the Constitutive View could deal with cases of empathizing with physical pains such as toothaches and migraines.

I wish to suggest that the implausibility of the Constitutive View has been obscured by the frequent characterization of empathy as empathic distress (see, e.g., Hoffman above). On this characterization of empathy, the term 'distress' is taken to be an apt description of both the empathic affect and of the affect felt by the suffering person. As Maibom puts it, "the term 'distress' is commonly taken to cover any of the individual emotions a person in need may feel" (Maibom, 2017).

If one characterizes empathy in this way, then one might reason as follows (i) whenever someone is suffering, they feel emotional distress, (ii) empathy with a suffering person thus involves feeling distress for that person and (iii) this distress would both by sympathetic (felt for the other) and empathic (a match of the other person's distress). This characterization of empathy is, I suggest, a gross simplification. People in need can feel a wide range of affective states, and thus empathy with such people can involve a wide range of affective states. Many

⁶¹ It may be noted that the way I am criticising the Constitutive View of the Empathy-Sympathy relationship is similar to the way that I critiqued the Intrinsic Motivation Explanation of empathy-induced altruism. This is not accidental. On the Constitutive View of the Empathy-Sympathy relationship, the Empathy-Sympathy Explanation of empathy-induced altruism becomes a version of the Intrinsic Motivation Explanation that I discussed above.

of those affective states do not count as distress and so do not count as sympathy, and thus the Constitutive View should be rejected.⁶²

4.4 The Causal View

In place of the Constitutive View of the empathy-sympathy relationship, we should instead adopt the *Causal View*.

The Causal View: When a person A empathizes with a suffering person (B), A's empathy with B causes A to feel additional sympathy for B.

This view is consistent with the empathic affect felt by A not being sympathy for B. On this view, there are two stages in the causal chain that gives rise to sympathy: the matching mechanism causes the empathic affect, and the empathic affect causes the sympathy. The Causal View leaves us with a new explanatory question: how does the empathic affect cause an increase in sympathy?

One might think that the answer to this question lies in the epistemic power of empathy. Perhaps it is by teaching us something about the other person's situation that empathy increases the sympathy that we feel for them. This view could draw on the idea, discussed above, that empathy is a way of finding out about the affective states of other people. This view also fits with Batson's model of sympathy, on which sympathy arises when we find out that someone we care

⁶² One might object here that the term 'distress' can be used in a broad way, to cover any kind of affective state that is constitutive of suffering. If one used the term 'distress' in that way, then I would concede that empathising with a suffering person would always involve feeling distress. I would deny, however, that feeling distress for another person entailed feeling sympathy towards them. Feeling depressed, guilty or exhausted for another person (if it makes sense to say that we can feel such affects for another person) might count as distress on this broad definition of distress, but it would not count as feeling sympathy for them.

about is in need (Batson, 2011: Ch. 2).⁶³ We could call this view the *Standard Epistemic Explanation**, now applied to the Empathy-Sympathy Hypothesis (whereas it was previously applied directly to explaining the Empathy-Altruism Hypothesis).

This revised version of the Standard Epistemic Explanation runs into the same problem as its counterpart. To be specific, it cannot explain how empathy can increase sympathy in a situation where the empathizer already knew what the other person was feeling. Recall the example of Steven and Oscar. If the Empathy-Sympathy Hypothesis is right, then when Steven feels empathy for Oscar, his sympathy towards Oscar will increase. If the Causal View is correct, this is because his empathy will cause an increase in his sympathy. The Standard Epistemic Explanation* of this increase is that it occurs because Steven's empathy for Oscar gives Steven new knowledge of what Oscar is feeling. This, however, cannot be right: Steven already knew what Oscar was feeling.

Nevertheless, I now go on to suggest, there is something that this explanation does get right: it correctly identifies that the key to empathy's motivating power lies in what it teaches us. The two Standard Epistemic Explanations say that empathy motivates us by giving us knowledge of what another person is feeling. I wish to finish this chapter by proposing that empathy motivates us by teaching us something beyond this. Specifically, I claim that empathy motivates us by giving us knowledge of what another person's feeling is like. I call this knowledge *phenomenal knowledge*. I refer to the idea that empathy

⁶³ For a similar account of the origins of sympathy, which takes inspiration from Aristotle, see Nussbaum (2001), especially section 6.2. It should be noted that (i) Nussbaum refers to sympathy as 'compassion', and (ii) Nussbaum also emphasizes the importance of seeing the other person's plight as being undeserved.

generates altruistic motivation by giving us such knowledge as the *Phenomenal Epistemic Explanation*.

5. The Phenomenal Epistemic Explanation

So far, this chapter has sought to show that there is a question about how empathy for a suffering person gives rise to altruistic motivation to help that person. I now outline how I think this question should be answered. The subsequent chapters of this thesis will help to fill out this outline.

Last year, I presented the example of Steven and Oscar to a group of high school students. They agreed that Steven's empathy for Oscar would increase Steven's altruistic motivation to help Oscar, and I pressed them to explain how Steven's empathy would do this. Most of them were drawn to the idea that Steven's empathy for Oscar was teaching him something about what Oscar was going through. I then pointed out that Steven already knew what Oscar was feeling, and asked: what more is there for Steven to learn? One student replied, 'Through his empathy for Oscar, Steven will know what it is like for Oscar to feel depressed'.

This is precisely the line of thought that I had been developing for this thesis, and it is how I think the Explanatory Question is best answered. As noted above, I call this line of thought the Phenomenal Epistemic Explanation. Applied to the case of Steven and Oscar, this explanation runs as follows. Before empathizing with Oscar, Steven knew that Oscar was depressed, but he did not know what it was like for Oscar to be depressed. When Steven empathizes with

Oscar, Steven acquires this further knowledge, and it is this further knowledge that motivates him (altruistically) to help Oscar.

This explanation can be broken down into two parts. First, there is a claim about what empathy teaches us and, second, there is a claim about how this empathy-induced learning motivates us. I call the first claim the *Strong Phenomenal Epistemic Claim*.⁶⁴

The Strong Phenomenal Epistemic Claim: If a person B is feeling an affective state S, then if person A empathizes with person B with respect to S, A's empathy for B will give A knowledge of what it is like for B to feel S, and this is the only way that A can obtain such knowledge.

My defence of this claim takes place in the next chapter, where I shall draw on the discussion of phenomenal knowledge that has been stimulated by Frank Jackson's example of Mary, the scientist who comes to see colour for the first time (Jackson, 1986). Defending this claim will require developing a new definition of empathy. Accordingly, I shall argue that what is essential to empathy is for the empathizer to be introspectively aware of their empathic affective states, and for the empathizer to use that awareness as a way of understanding the feelings of the person that they are empathizing with.

⁶⁴ As I shall discuss at the start of the next chapter, this claim is strong because it says that empathy is the only source of phenomenal knowledge about the affective states of other people. A weak version of this claim would say only that empathy is one source of such phenomenal knowledge, without saying that it was the only source. The Phenomenal Epistemic Explanation of empathy-induced altruism could perhaps be made to work by using this weak version of the claim (instead of the strong version). However, the Phenomenal Epistemic Explanation would then fail to work in cases in which, prior to empathizing with the other person, the empathizer already had phenomenal knowledge of what the other person was feeling. The Strong Phenomenal Epistemic Claim makes such cases impossible.

In chapters 4 and 5, I then go on to explore the significance of phenomenal knowledge for moral deliberation. In chapter 4, I argue that if we know what a person's suffering is like (which we can only do through empathizing with them), then we will know how bad their suffering is for them. In chapter 5, I extend this idea, and argue that if we know what a person's suffering is like, we will know the strength of reason that their suffering gives us to help them. Although the primary aim of these chapters is to vindicate the role of empathy in moral deliberation, they also serve to strengthen the second part of the Phenomenal Epistemic Explanation. In other words, they help to give credibility to the idea that acquiring phenomenal knowledge about another person's suffering (knowing what their suffering is like) will motivate us to help them. At the end of chapter 5, I will show how the Phenomenal Epistemic Explanation of empathy-induced altruism can be developed along these lines.

Chapter 3: Empathy and Phenomenal Knowledge

1. Overview

This chapter argues that empathy plays a distinctive role in inter-personal understanding. More specifically, it argues that whilst there are many ways to know which affective states (emotions, moods and physical sensations) another person is feeling, it is only through empathy that we can know *what it is like* for that person to feel such affective states. I call the former kind of knowledge *attributive knowledge*, because it involves attributing affective states to other people. I call the latter kind of knowledge *phenomenal knowledge*, because it is knowledge about the phenomenal quality of the affective states of other people. Using this terminology, my claim is that empathy is a source, and the only source, of phenomenal knowledge about the affective states of other people.⁶⁵

I shall refer to this claim as the *Strong Phenomenal Epistemic Claim*. The full version of this claim is as follows.

The Strong Phenomenal Epistemic Claim: If a person B is feeling an affective state S, then if person A empathizes with person B with respect to S, A's empathy for B will give A knowledge of what it is like for B to feel S, and this is the only way that A can obtain such knowledge.

To illustrate this claim, let us return to the example of Steven and Oscar from the previous chapter. Oscar is depressed, and Steven is a friend of Oscar's who is

⁶⁵ I do not mean to say that empathy is the only source of phenomenal knowledge. Rather, I am saying that it is the only source of phenomenal knowledge with respect to the affective states of other people. We can have phenomenal knowledge of our own affective states through introspection. Moreover, although empathy is the way through which we obtain phenomenal knowledge of the affective states of other people, we can also have phenomenal knowledge of the perceptual states of other people, and I am not claiming that this is obtained through empathy (at the end of section 3 below, I claim that it is obtained through a near-relation of empathy).

trying to understand what Oscar is going through. Steven has already learned that Oscar is feeling depressed by listening to Oscar describe his feelings (thus acquiring attributive knowledge). Steven could perhaps also have learnt this by making an inference from Oscar's reclusive behaviour, or from his downcast facial expressions. According to the Strong Phenomenal Epistemic Claim, however, it is only through empathizing with Oscar that Steven can know what it is like for Oscar to feel depressed (thus acquiring phenomenal knowledge). This is consistent with the analysis of this example that I gave in the previous chapter.

The Strong Phenomenal Epistemic Claim is strong because it says that empathy is the only route to phenomenal knowledge of the affective states of other people. A weak version of this claim would say that empathy is one source of such knowledge, but would not say that it is the only source. Consequently, the weak version of this claim would not vindicate empathy as playing a distinctive role in interpersonal understanding.

The Strong Phenomenal Epistemic Claim can be challenged by any view which says that phenomenal knowledge of the affective states of others can arise without the use of empathy. One such view, which commands intuitive support, is given by what I call the *Sufficiency of Acquaintance Claim*.⁶⁶ This claim says that the crucial factor in determining whether I know what someone else's experience is like is whether I have had a similar experience myself.

The Sufficiency of Acquaintance Claim: If a person B is feeling an affective state S, then if person A has had at least one experience of an affective

⁶⁶ For a brief statement of this view, see Matravers (2011, p. 28-30).

state that is similar to S, and if A knows that B is feeling S, then A knows what it is like for B to feel S.⁶⁷

Once again, let us illustrate this view by returning to the example of Steven and Oscar. As I set up that example in the previous chapter, Steven had experienced depression before. He also knows that Oscar is depressed. Although he has not yet empathized with Oscar, the Sufficiency of Acquaintance Claim says that Steven knows what it is like for Oscar to feel depressed, because Steven has also been through depression. We might imagine Steven saying to Oscar, 'I know what it's like because I had a really difficult few months a couple of years ago'.

The Sufficiency of Acquaintance Claim and the Strong Phenomenal Epistemic Claim thus give conflicting interpretations of the case of Steven and Oscar. The former says that, prior to empathizing with Oscar, Steven knows what Oscar's depression is like. The latter denies this. According to the Strong Phenomenal Epistemic claim, it is only once Steven empathizes with Oscar (perhaps by utilizing his own past experiences of depression) that Steven knows what Oscar's depression is like. The Strong Phenomenal Epistemic Claim says that Steven's own experiences of depression might help Steven know what Oscar's experience of depression is like, but it says that it will only do this if Steven utilizes those experiences to empathize with Oscar. If he does not do this,

⁶⁷ There is a question here about how similar the affective states of A and B need to be, in order for it to be the case that A knows what B's affective state is like. A first approximation is that they must be the same kind of affective state, and that they must be of approximately the same intensity. Thus, in the example above, the Sufficiency of Acquaintance View entails that Steven knows what Oscar's experience of depression is like just if Steven's earlier experiences of depression were of a similar intensity (had Steven's depression been much milder than that which Oscar is presently experiencing, we would say that he did not know what Oscar's depression was like). It should be noted that the Sufficiency of Acquaintance Claim does not say that person A and person B need to have been through the same kind of external circumstances. Rather, it says only that they have felt the same kind of affective state. Accordingly, if B is feeling grief about the death of a sibling, and A has felt a similar grief about the death of a close friend, A can still be said to know what B's grief is like.

the Strong Phenomenal Epistemic Claim says, he will not know what Oscar's experience of depression is like.

My position in this chapter is that the Strong Phenomenal Epistemic Claim is correct. My primary defence of this position will come in three stages. First, in section 3, I offer an account of empathy as *experiential understanding* of the affective states of other people. So far, all that I have said about empathy is that it is a form of affective matching. There are, however, some forms of affective matching that do not yield phenomenal knowledge. Consequently, in order to defend the Strong Phenomenal Epistemic Claim, I shall give a more specific definition of empathy. Secondly, in section 4, I show that there are two kinds of everyday case which undermine the Sufficiency of Acquaintance Claim and which lend support to the Strong Phenomenal Epistemic Claim. Thirdly, in section 5, I give an argument in favour of the Strong Phenomenal Epistemic Claim.

In section 6, I respond to two possible objections that might be made to my defence of the Strong Phenomenal Epistemic Claim. The first objection is that I have not done enough to show that the Sufficiency of Acquaintance Claim is false. Given that the Sufficiency of Acquaintance Claim is incompatible with the Strong Phenomenal Epistemic Claim, the objection would continue, this means that I have not done enough to show that the Strong Phenomenal Epistemic Claim is true. To meet this objection, I show that the Sufficiency of Acquaintance Claim, when formulated in its most plausible version, is consistent with the Strong Phenomenal Epistemic Claim because it employs a different conception of knowledge. The second objection is that I have not done enough to show that it is correct to define empathy as the experiential understanding of the affective states

of other people. I respond to this objection by offering an argument in favour of my definition of empathy.

Before beginning my defence of the Strong Phenomenal Epistemic Claim, it will be helpful to complete a preliminary task. In section 2, I contextualise the Strong Phenomenal Epistemic Claim against the backdrop of philosophical debates that have been taking place within the philosophy of mind and the philosophy of consciousness. I suggest that it has not received as much attention as it deserves in either of those debates.

Before engaging with this task, it is worth recapping how the Strong Phenomenal Epistemic Claim fits in with the discussion of empathy and altruism that I offered in the previous chapter. At the end of that chapter, I suggested that the most promising explanation of empathy-induced altruism was the Phenomenal Epistemic Explanation. That explanation says that empathy for a suffering person increases our altruistic motivation to help that person by giving us new phenomenal knowledge of what that person's suffering is like. The first part of this explanation is thus the claim that empathy is the only source of phenomenal knowledge about the affective states of other people. In other words, it is the Strong Phenomenal Epistemic Claim that I am defending here.

The overarching goal of this thesis is to show that empathy has an important role to play as an input into moral deliberation. This chapter lays a crucial foundation for that defence by showing that empathy makes a distinctive contribution to interpersonal understanding: empathy alone is a source of phenomenal knowledge about the affective states of other people. In subsequent chapters (4 and 5) I progress to the task of demonstrating the importance of

phenomenal knowledge to moral deliberation. In short, the idea is this: where another person is suffering, then by knowing what their suffering is like you can know how bad it is for them and you can know the strength of your reason to help them.

2. Literature Review

The Strong Phenomenal Epistemic Claim has not received the attention that it merits. In combining the notions of empathy and phenomenal knowledge, it has fallen into the gap between two debates, each of which has considered one notion, but not the other. On the one hand, the *Mindreading Debate* has considered empathy to be a source of knowledge but has not explored the possibility that the knowledge given to us by empathy is phenomenal knowledge. On the other hand, the *Knowledge Argument Debate* has considered phenomenal knowledge but has not explored the possibility that it is empathy that gives rise to that knowledge.

2.1 Empathy and the Mindreading Debate

The Mindreading Debate has sought to explain our everyday ‘folk psychological’ ability to know which mental states other people are experiencing.⁶⁸ Using the terminology that I introduced above, we can say that it seeks to explain how we acquire attributive knowledge. Although this debate

⁶⁸ In addition to the defences of simulation theory that I cite below, influential (and contrasting) contributions to this debate include Gopnik & Wellman (1992), Dennett (1987) and Gallagher (2001).

initially focused on how we attribute intentional mental states (such as beliefs and desires) to other people, there has more recently been a significant amount of discussion about how we attribute affective states to other people. One influential approach argues that simulation is an important mechanism through which we do this (Harris, 1992; Jackson, Meltzoff and Decety, 2005; Goldman, 2006).

Given the close relationship between simulation and empathy, this approach lends support to the idea that empathy is a source of attributive knowledge about the affective states of other people.⁶⁹ This idea has been taken up in recent defences of the moral significance of empathy: if such knowledge is an important input into moral deliberation, and if empathy gives us such knowledge, then it follows that empathy is an important input into moral deliberation (Masto, 2015; Song, 2015). However, even if empathy is a source of attributive knowledge, this does not show that empathy has a distinctive epistemic importance, because there are other sources of attributive knowledge. We can find out what others are feeling by asking them (thereby listening to their testimony), or by making inferences from their behaviour, utterances and facial expressions. This raises a question: why should we empathize with others, when there are alternative (perhaps easier) ways of finding out what they are feeling?

Maibom (2007) suggests that what distinguishes empathy from testimony and inference is its motivational power: unlike testimony and inference, empathy involves feeling what the other person feels, and thus it necessarily motivates us to help them. As I discussed in the previous chapter, I agree with Maibom's view

⁶⁹ Roughly speaking, the term 'simulation' has here been used to refer to the process of imagining being in another person's situation and observing which psychological states one thereby comes to be in. 'Simulation' thus bears a close connection to what I am calling 'perspective-taking'. I have identified perspective-taking as an important mechanism through which empathy can arise.

that empathy gives rise to altruistic motivation. My objection to her approach is that it is mistaken about how empathy motivates us. Maibom thinks that empathic affective states motivate us directly, whereas I think that they motivate us by giving us phenomenal knowledge. On my view, what is special about empathy is that it, and it alone, can give us phenomenal knowledge about the affective states of other people.

2.2 Empathy and the Knowledge Argument Debate

Within the philosophy of consciousness, the Knowledge Argument Debate has offered extensive discussions of phenomenal knowledge. This debate was inspired by a thought experiment that was originally penned by Frank Jackson (Jackson, 1982).⁷⁰ Jackson asks us to consider the (hypothetical) case of Mary, a scientist who, by being confined to a black and white room for the entirety of her life, has never seen colour.⁷¹ Nevertheless, Mary knows everything that there is to know about the physics of colour perception: Mary knows about the nature of light-waves, about how and why different objects reflect light-waves differently, and about how light-waves are processed by our eyes and brains to generate our visual perceptions.

Mary is capable of knowing that other people are having mental states of colour perception (on my terminology, she can have the relevant attributive knowledge). For example, she can know that a colleague John is seeing a blue

⁷⁰ Nagel's earlier discussion of the challenges of understanding the mental life of a bat has also been influential (Nagel, 1974). For a useful anthology of responses to Jackson's article (including Jackson's own revision of his view), see Ludlow, Nagasawa and Stoljar (2004).

⁷¹ Knut Nordby was a real life case of a colour vision specialist who could not see colours. See Nordby (2007) for an interesting (and autobiographical) philosophical reflection on his condition.

sky. In fact, she can know a lot more than this, because she knows everything that there is to know about the physics of colour perception. What Mary cannot know, according to Jackson, is what it is like to have such mental states.⁷² For example, she cannot know what it is like for John to see a blue sky. On the terminology that I am using here, we can say that Mary cannot have phenomenal knowledge about another person's mental states of colour perception.

Mary's epistemic position changes when she sees colour for the first time. When she is released from her black-and-white room and encounters the coloured entities in the outside world, she will see (amongst other things) the blue sky. Upon doing this, she could exclaim, 'Ah, this is what it is like to see the blue sky! Now I can know what John has been experiencing all along!' Now that Mary has acquired her own experiences of seeing blue, she can know what it is like for others to see blue.

Jackson introduced the example of Mary as part of his rebuttal of Physicalism (the view that all facts are physical facts). Jackson reasoned that if Mary previously knew all of the physical facts about colour perception, and then learnt a new fact about colour perception (a phenomenal fact), it follows that this new fact cannot be a physical fact. In which case, not all facts are physical facts and Physicalism is false. In response to Jackson's argument, a complex debate

⁷² Although most philosophers agree with this claim, Dennett denies it (Dennett, 2002). Dennett thinks that if Mary really did know all of the physical facts prior to her release, she could know what it is like to see colour. Dennett's view may thus seem to undermine the Strong Phenomenal Epistemic Claim that I defend here, as he could say that an expert in psychology (who was omniscient about the physical facts associated with feeling affective states) could know what it is like to feel an affective state without ever having to feel it herself, or without ever having to empathize with another person. However, given that such omniscience is hard to come by, Dennett might still be persuaded to adopt a weakened version of the Strong Phenomenal Epistemic Claim which said that, given practical limitations about how much knowledge we can acquire without empathizing with others, empathy is the only way in which we can acquire phenomenal knowledge about the affective states of other people.

has emerged as to the nature of phenomenal knowledge. Although most philosophers agree that Mary acquires new knowledge upon her release, many have developed accounts of phenomenal knowledge on which this learning is consistent with Physicalism.⁷³

When considering the conditions under which phenomenal knowledge arises, a common idea within this debate is that acquaintance is a necessary condition for phenomenal knowledge. Thus, until Mary has seen blue, she cannot know what it is like to see blue. More generally, it can be said that one cannot know what it is like to have a certain sort of mental state until one has had a mental state of that sort.

As I will go on to explore, the idea that acquaintance is necessary for phenomenal knowledge raises a question: how do we acquire phenomenal knowledge, and what is the role of acquaintance in this process such that it is an essential part of it? Although the literature in the Knowledge Argument Debate has engaged with this question, it has not utilized the notion of empathy to answer it. One explanation for this is that the debate has focused on cases of having phenomenal knowledge with respect to the perceptual mental states of other people (such as seeing a blue sky) whereas empathy involves relating to the affective mental states of other people.

⁷³ The two main strategies are: (i) to argue that phenomenal knowledge is not propositional, and thus that acquiring it does not involve learning a new fact (Lewis, 1988; Nemirow, 1990; Conee, 1994), and (ii) to argue that the phenomenal knowledge acquired by Mary, though propositional, is of a fact that she already knew through a different mode of representation (Horgan, 1984; Churchland, 1985; Tye, 1986; Papineau, 2002). As far as possible, I wish to avoid the question of what phenomenal knowledge is, and focus on identifying the conditions under which it arises. That said, I think that my rejection of the Sufficiency of Acquaintance View entails a rejection of the first strategy identified above, and thus I am arguing here for the idea that phenomenal knowledge is propositional.

2.3 Supporters of the Phenomenal Epistemic Claim

Although the Strong Phenomenal Epistemic Claim has not received the attention that it merits, it has been endorsed by a small number of contemporary philosophers (Coplan, 2011; Masto, 2015; Ravenscroft, 2017; Smith, 2017).⁷⁴ Instead of arguing for the Strong Phenomenal Epistemic Claim, however, these philosophers have simply asserted it. Although they have been careful to define empathy in a way that supports the Strong Phenomenal Epistemic Claim (as I shall do in section 3 below), they have not gone on to provide an argument for that claim. Moreover, they have not defended it against the challenge posed by the Sufficiency of Acquaintance Claim.

One philosopher who has provided an argument that might be used to support the Strong Phenomenal Epistemic Claim is Biggs (2009). Although Biggs does not mention empathy, he uses the term ‘phenomenal simulation’ to refer to what I will be calling ‘empathy’. In comparison with alternative ways of attributing mental states to other people, he claims that phenomenal simulation provides a deeper form of understanding. His argument for this claim is that only phenomenal simulation employs phenomenal concepts, and that phenomenal concepts deepen our understanding of the mental states of other people. A phenomenal concept is a concept that can only be deployed by a person if that person instantiates the relevant phenomenal property. For example, to use the

⁷⁴ Each of these philosophers states the claim in a different way. Coplan writes that empathy is a ‘unique kind of understanding through which we can experience what it is like to be another person.’ (Coplan, 2011: p. 6). Masto suggests that what is special about empathy is that ‘in experiencing the affect of the target, the empathizer comes to know how it feels to feel like that’ (Masto, 2015: p. 85). Using similar language, Smith claims that: ‘there is an epistemic function of empathy that is, I suggest, distinctive of it. Empathy provides us with knowledge of *how others feel*.’ (Smith, 2017: p. 712).

phenomenal concept of red, one must be accessing a present moment experience of redness. Because Biggs does not argue specifically for the claim that phenomenal simulation (and thus, on my terminology, empathy) is the only source of phenomenal knowledge, I shall not reproduce his argument at length here. However, I shall draw on it below when I offer my own argument for the Strong Phenomenal Epistemic Claim.

Having contextualised the Strong Phenomenal Epistemic Claim against the backdrop of recent philosophical debates, I now begin my defence of it. My first task in that defence is to introduce my definition of empathy as experiential understanding.

3. Empathy as Experiential Understanding

So far in this thesis, all that I have said about the definition of empathy is that empathy is a kind of affective matching. In other words, I have said that there is an Affective Match Condition on empathy.

The Affective Match Condition: For it to be the case that person A is empathizing with person B with respect to B's feeling S (an affective state), it must be that A feels S*, where S* is similar to S.

In order to defend the Strong Phenomenal Epistemic Claim, I need to offer a more specific definition of empathy. That is the task that I turn to in this section.

I define empathy to be the *experiential understanding* of another person in relation to an affective state that they are experiencing.⁷⁵ On this definition, empathy can be said to be taking place just when three conditions are met.

First, the affective match condition (restated above) must be met. My interpretation of the similarity requirement in this condition is as follows. For it to be the case that S* can be said to be similar to S, S* and S need to be affective states of the same type (e.g. both anger), and of approximately the same intensity. Where the affective states are emotions, there is a further question about whether they need to have the same objects. My view is that they do not need to. Thus, I could empathize with a friend's frustration about politics by feeling a similar (and similarly intense) kind of frustration about sport.

On my approach, having an affective state that is similar to the one that is felt by the other person puts you in a position to empathize with what the other person is feeling. It does not, however, guarantee that you will empathize with them. For such empathy to arise, a second and third condition must be met.

The second condition assumes that the Affective Match Condition has been met and adds a requirement about how the empathizer relates to their empathic affect. In other words, it adds a requirement about how person A relates to S*. I call this requirement the *Introspective Awareness Condition*.

⁷⁵ I have borrowed the term 'experiential understanding' from Coplan (2011) because it captures what is distinctive about empathy as a way of learning about the mental life of other people. Although my definition of empathy is similar to that given by Coplan, it differs from it in two important ways: (i) on Coplan's definition, empathy can only arise through perspective-taking, whereas I allow that it can also arise through affective contagion or through one's own lived experience, (ii) Coplan does not explicitly make introspective awareness of the empathic affect a necessary feature of empathy, whereas I do.

The Introspective Awareness Condition: Person A is introspectively aware of S*.

To be introspectively aware of an affective state is to attend to how it feels.⁷⁶ In other words, it is to put one's attention on its phenomenology (that is, on what it is like). The phenomenology of an affective state might include: body sensations (e.g. a burning in one's chest during anger), the experience of related thought patterns (e.g. interpretations of why another person deserves punishment), or a felt quality of the affective state that is not reducible to either of the aforementioned things.⁷⁷ As I shall argue in the next chapter, the phenomenology of an affective state will also be influenced by the extent to which we dislike that affective state.

The Introspective Awareness Condition is important because it is possible to feel an affective state without being introspectively aware of it. One might instead, for example, be attending to a practical task or to another affective state. To illustrate, if I am stressed about work, I may spend the whole day thinking about what I need to do, and never once attend to how it feels to be stressed. If, at the end of the day, a friend asks me how I am feeling, my subsequent introspection may bring a sense of revelation. I may exclaim, 'Wow, I did not realize how stressed I have been!' Even if I had already realized that I was stressed, my introspection could still add further knowledge about how it feels (and about how bad it feels) to be stressed.

⁷⁶ In a fascinating and relevant discussion, Goldie refers to this introspective awareness as 'reflective awareness' or 'reflective consciousness' (Goldie, 2002: pp. 62-68). Goldie clarifies that one can have an emotion, and know that one is having an emotion, without having reflective awareness of that emotion.

⁷⁷ For a similar, and more extended, account of the phenomenology of emotion, see Goldie (2002: pp. 52-62). If the affective state is a physical pain or pleasure, then its phenomenology will simply be that of a body sensation.

We are particularly liable to avoid introspecting painful affective states, because introspecting such states can seem to bring us closer to the pain. If I find myself in emotional pain, I am instinctively drawn either to distract myself (perhaps by watching television, by getting some unrelated work done, or by eating something), or to strategize about how to eliminate the pain (for example, if I am feeling lonely, I may start planning ways to meet up with friends). Neither of these responses involve attending to how the painful affective state feels. If one trains in the Buddhist-inspired meditation techniques that are now being offered as part of mindfulness programs, it is common to be asked to let go of these responses, so that we can spend more time introspecting difficult affective states.⁷⁸

My experience has also been that strong emotions tend to make me attend to the object of the emotion, rather than to the phenomenology of the emotion itself. For example, if I am feeling angry at a friend, I frequently become lost in thoughts about what they have done, and about how I should respond to their actions. I rarely stop to notice how the anger feels. Similarly, if I am feeling stressed about work, my focus will be on how to get the work done, rather than on introspecting the stress.

Just as one can fail to introspect one's own affective states, one can also match the affective state of another person without introspecting one's corresponding empathic affect. For example, I might pick up the fear of another person through affective contagion, but fail to be introspectively aware of this empathic fear because I am distracted by a meal that I am eating, or by my own

⁷⁸ See, for example, 'Chapter 9: Turning Towards Difficulties' in Williams and Penman (2011).

affective states. Moreover, feeling empathic fear could lead me to be lost in thoughts relating to what the fear is about (e.g. an upcoming deadline, or a loved one's health) rather than in noticing the phenomenology of the fear itself. As with our own affective states, if the empathic affective state is very painful or powerful, then it will be difficult to introspectively attend to it.⁷⁹

The third condition on empathy assumes that the first two conditions have been met, and adds a final requirement. I call this the *Phenomenal Attribution Condition*.

The Phenomenal Attribution Condition: Person A has a conscious and justified belief that person B is feeling an affective state that is 'like this', where 'this' refers to the phenomenal quality of S*.

This condition is important because the previous two conditions could be met without person A having any awareness that they are empathizing with person B. For example, I could feel a fear that is similar to my friend's fear, and introspect on my fear, without giving any thought to the fact that my friend is similarly afraid. The Phenomenal Attribution Condition adds the additional requirement that I have the thought, 'My friend is feeling an affective state that feels just like this (my fear)'.

In requiring that the empathizer has a conscious belief that B is feeling something 'like this', I mean that the empathizer is actively attending to that belief. In other words, I mean that it is something that they are consciously aware of. In

⁷⁹ The difficulty of empathizing with the pain of others has been expressed forcefully by Simone Weil: 'The capacity to give one's attention to a sufferer is a very rare and difficult thing; it is almost a miracle; it is a miracle. Nearly all those who think they have this capacity do not possess it. Warmth of heart, impulsiveness, pity are not enough/' (Weil, 1959/2009: p. 64)

requiring that the empathizer's belief is justified, I mean to rule out the case in which the would-be empathizer believes without justification that they are feeling what someone else is feeling and, through complete chance, happens to be correct. What would count as a justification here? First, the empathizer could have matched the other person's affective state through a reliable matching mechanism (such as perspective-taking or affective contagion). Alternatively, they could have made a justified inference (e.g. on the basis of the other person's situation, or their observable behaviour) about what the other person is feeling.⁸⁰

Unlike many definitions of empathy, my approach does not place any restrictions on how the empathic affective state (S^*) comes about. My approach allows that the empathic affective state could come about through affective contagion or through consciously mediated perspective-taking. It also allows that it could come about without the involvement of either of these matching mechanisms. What I have in mind here is the possibility that the empathic affective state could come about through the empathizer's own lived experience. This is an important possibility that has been overlooked within the contemporary debate.

As we go through our everyday lives, we will encounter a range of different affective states in response to the ups and downs that we go through. We might be stressed about a deadline, grateful for a friendship, frustrated about a

⁸⁰ In this thesis, I will not be engaging with sceptical scenarios concerning our ability to know what the experiences of other people are like. Two such scenarios are: (i) that other people are philosophical zombies that have no experiences whatsoever, even if they continue to act like people who do (Chalmers, 1996), and (ii) that other people can experience mental states very differently to the way that we experience them (Shoemaker, 1982 and Block, 1990). I do not think that empathy can be used to rule out such scenarios, and I do not think that a person is required to rule them out in order to be able to empathize with another person. One might therefore restate the Strong Phenomenal Epistemic Claim in the following way: assuming that other people experience affective states in the same way that we do, empathy is a way (and the only way) of coming to know what their affective states feel like.

colleague, or excited about an upcoming holiday. If, at any point, you know that another person is experiencing an affective state that is similar to the one that you are experiencing, then you can use your affective state to empathize with that person. All that is required is for you to introspect your affective state and to (justifiably) think that the other person is feeling something 'like this'. In a similar way, one can also use one's memories of past affective states to empathize with what another person is presently experiencing.

Sometimes, I find that my own life experience can give me insights about what another person has been going through. For instance, if I have been experiencing one side of a situation involving another person (e.g. trying to schedule rehearsals for a theatre show), I may be focused on what I am feeling (the stress of trying to co-ordinate people) and lose sight of what they are feeling (the frustration of being asked to attend rehearsals, in the midst of a busy period). If, in the future, I am on the other side of the situation, then I may find myself experiencing the same feeling (frustration) that I had previously been inducing in the other person. This can lead to an 'ah hah' moment where I realize what the other person had been experiencing all along. I can use my new experience to empathize with what they were going through.

One can also deliberately cultivate new life experiences in order to empathize with other people. In a chapter called 'Seek Experiential Adventures', Roman Krznaric gives a series of examples of people who have done this as a way of engaging with oppressed and underprivileged groups (Krznaric, 2014: ch. 3). Krznaric begins by discussing the British social reformer Beatrice Webb:

In 1887, her research took a more experiential turn when she dressed up in a bedraggled skirt and buttonless boots in search of work as a

seamstress in an East End textile sweatshop...Crammed in with thirty other women and girls, she did piecework on twelve-hour shifts.. (ibid.: pp. 74-75)

Some thirty years later, George Orwell did something similar. He spent time living, sleeping and eating alongside homeless people in the streets of East London. In his own words, his intention was as follows:

I wanted to submerge myself, to get right down among the oppressed... to be one of them and on their side against the tyrants'. (Orwell, 1962: p. 130)

In the 1980s, German investigative journalist Günter Wallraf spent months pretending to be a Turkish immigrant and working in the kind of jobs that such immigrants were able to get access to in Germany:

I experienced conditions which are usually only described in history books about the nineteenth century. The work was dirty, crushing and drained one's last reserves; but worse was the humiliation that I had to bear and the contempt in which I was held. (Wallraf, 1988: p. 177)

Wallraf's experiment in empathy turned out to have long-lasting effects: he was scarred with chronic bronchitis following his work shovelling coke dust without a protective mask.

In cases like these, empathy is facilitated in so far as the empathizer comes to have affective states that are similar to those that are felt by the person that the empathizer is trying to empathize with.⁸¹ These are cases of literal perspective-taking. That is, they are cases of entering into an external situation that is similar to that which is experienced by the other person. By literally taking

⁸¹ Of course, cases like these may fail to generate genuine empathy. First, a simulated experience may fall far short of the actual one (living as a sweatshop worker for a few days would not feel the same as doing that job for a lifetime). Secondly, the would-be empathizer might become so overwhelmed by their experience that they neither introspect their affective states, nor use them to understand the experiences of other people. In other words, the would-be empathizer might become so lost in their own suffering that they lose sight of their intention of empathizing with what other people are going through.

up the other person's perspective, the empathizer can come to feel as they feel without needing to imagine anything. Such cases have been overlooked in a literature that focuses on imaginative perspective-taking.

The preceding paragraphs have been focused on showing how, on my definition of empathy, one can empathize with other people by drawing on one's own lived experience. In other words, my point is, empathy need not necessarily arise through perspective-taking or through affective contagion.

As a final point in this section, it is important to note that although my definition of empathy involves experiential understanding with respect to the affective states of other people, one can also have experiential understanding with respect to any mental state that has a phenomenal quality. For example, one could have experiential understanding with respect to the perceptual states of another person. Such experiential understanding would involve having a similar perceptual state, introspecting on that perceptual state, and thinking that the other person is experiencing something 'like this'. We could imagine Mary doing this upon her release. Looking up at the blue sky, she could introspect that perceptual state and exclaim, 'John has been experiencing something like this for all these years!'. I refer to this kind of experiential understanding as *perceptual-empathy*, and reserve the term 'empathy' specifically for cases of experiential understanding with respect to the affective states of other people.

4. Two Kinds of Everyday Example

Having clarified how I propose to use the term 'empathy', I now move on to the task of defending the Strong Phenomenal Epistemic Claim. My first step will

be to outline two kinds of everyday case which undermine the Sufficiency of Acquaintance Claim and support the Strong Phenomenal Epistemic Claim.

4.1 Acquainted but phenomenally ignorant

To begin, recall that the Sufficiency of Acquaintance Claim says that if I know which affective state another person is experiencing, and if I have had at least one experience of a similar affective state, then I know what their experience is like. I suggest that this claim is shown to be false by everyday situations in which we know which affective state another person is experiencing, and have experienced it ourselves, yet in which we fail to know what the other person's experience is like. To the extent that the Sufficiency of Acquaintance Claim is the main rival to the Strong Phenomenal Epistemic Claim, these cases also provide (indirect) support for the Strong Phenomenal Epistemic Claim.

Here is an example of the kind that I have in mind, drawn from my own experience.

The Difficulty of Shared Grief: Several of my friends recently have had one of their parents die. At risk of making a gross simplification, let us say that there is a characteristic kind of grief that people experience in these circumstances. My own father died, so I have personal experience of such grief. In other words, we can say that I am acquainted with grief. However, my father died twenty years ago, and I find it difficult to access my memory of these feelings (perhaps it is more accurate to say that I find it difficult to consciously access the feelings that still persist). Perhaps the limiting factor here is that I do not want to consciously access these feelings,

because they are painful, and potentially overwhelming. For these reasons, I have not empathized with my friends. Moreover, I judge myself to be failing to adequately understand them, even though I have experienced grief, and I know that they are grieving. My interpretation of this failure of understanding is that it consists in not knowing what it is like for my friends to grieve.⁸²

Of course, one might reply to this example by maintaining that I am in a better position to understand what my friends are going through than someone who has not experienced grief. This may be true. Having my own experiences of grief will help me to empathize with my friends. My contention is, however, that unless I actually empathize with my friends, I will not know what their grief is like.

4.2 Empathic epiphanies

The second type of case that I wish to consider is what I call ‘empathic epiphany cases’. These cases have two stages. In the first stage, we are acquainted with what another person is going through, and know which affective state they are experiencing, but we are not empathizing with them. In the second stage, we empathize with the other person and have an intuition that we are learning something new. I suggest that this intuition is correct and that in the second stage we acquire new knowledge of what the other person’s affective state is like. For us to be able to acquire this knowledge, we must have previously

⁸² I also have a conflicting intuition that I do know what my friends’ grief is like, because I have experienced it myself. This intuition appears to support the Sufficiency of Acquaintance Claim. In section 6.1 below, I explain how this intuition can be reconciled with the Strong Phenomenal Epistemic Claim by distinguishing between two different kinds of knowledge.

lacked it. Accordingly, the first stage of this type of case is also an example of the ‘acquainted but phenomenally ignorant’ cases discussed above. What is distinctive about these cases is that our initial ignorance is demonstrated by the intuition that we learn something new when we empathize with the other person.⁸³

I now present an example of such a case. It is drawn from my own experience and it was the inspiration for the example of Steven and Oscar that I introduced in the previous chapter.

Lagged empathy for a depressed friend: Several years ago, I spent an hour listening to a friend who was feeling very low. He was self-aware, and was pro-active in describing his emotional state, so my impression was that I formed an accurate understanding of what he was feeling: some combination of lethargy, hopelessness, fear, and shame (I could have also added more detail on what these feelings were about, and how intense they were, and so on). I did not, however, empathize with my friend. Emotionally, I felt detached, and because of this I was not satisfied that I was really understanding my friend. Later in the day, I had a sudden and vivid memory of a time when I had experienced similar emotions. I felt what my friend was going through. I had a sense of revelation as I thought, ‘Oh, this is what my friend has been experiencing.... now I understand’.

My interpretation of this example is that the additional understanding that I acquired through empathizing with my friend consisted in knowing what my

⁸³ Biggs offers a similar kind of example in support of his argument that empathy (which he calls phenomenal simulation) has a distinctive epistemic importance (Biggs, 2009: pp. 650-653). However, instead of having one person go through two stages of understanding another person in pain, Biggs has two different people try to understand a person in pain (one of whom empathizes, and one does not). The weakness of this approach is that it requires the reader to imagine the extent to which each of those people understands the person in pain. On my approach, we can simply reflect on our own experiences of empathizing with another person who is in pain.

friend's experience was like. In other words, my empathy for him gave me the phenomenal knowledge that I had previously lacked. If this is right, then the existence of this type of case shows that the Sufficiency of Acquaintance View is false (because I initially failed to know what my friend's depression was like, even though I was acquainted with depression). This example also supports the idea that empathy is a source of phenomenal knowledge.

Do examples such as this prove that empathy is the only source of phenomenal knowledge? Not as it stands. One could argue that I could have also acquired knowledge of what my friend's depression was like without using empathy. Perhaps I could have extensively researched the neuroscience and physiology of depression, and used cutting-edge technology to learn exactly what was going on in my friend's body and brain whilst he was depressed. If I had done these things, one might think, I could have acquired knowledge of what my friend's depression was like, and thus when I empathized with him, I would not have learnt anything new.

In response to this line of thought, I offer three considerations. First, it might be that the kind of research identified above would give me phenomenal knowledge by helping me to empathize with my friend. For instance, perhaps reading about what depression does to one's body would help me vividly imagine an experience of being depressed. The Strong Phenomenal Epistemic Claim is compatible with psychological research giving rise to phenomenal knowledge in this way. The scenario which threatens the Strong Phenomenal Epistemic Claim is the one in which such research fails to stimulate empathy for my friend, but still generates knowledge of what his experience is like. Secondly, I suggest that most philosophers would be sceptical that such a scenario could take place, for the

same reasons that they are sceptical of the claim that Mary can know what it is like to see a colour before she has been released from her black and white confines. In the next section, I will connect my defence of the Strong Phenomenal Epistemic Claim to the discussion of the Mary case. Thirdly, even if I conceded that an omniscient psychologist might be able to acquire phenomenal knowledge without empathy, I could still maintain that such omniscience is unobtainable in practice and thus that, in practice, empathy is still the only way we have of acquiring phenomenal knowledge about the affective states of other people.⁸⁴

My goal in this chapter, however, is to defend an unmodified version of the Strong Phenomenal Epistemic Claim. In other words, I am arguing that even if we did not face practical limitations on our epistemic abilities, empathy would still be the only means that we had of acquiring phenomenal knowledge of the affective states of other people.

The purpose of this section has been to present two examples which undermine the Sufficiency of Acquaintance Claim and lend support to the Strong Phenomenal Epistemic Claim. I now give an argument for the latter claim.

5. An Argument for the Strong Phenomenal Epistemic Claim

In this section, I offer an argument for the Strong Phenomenal Epistemic Claim.⁸⁵ The first premise of this argument is a widespread view about the conditions under which phenomenal knowledge can arise.

⁸⁴ For a related discussion, see footnote 72 above.

⁸⁵ As noted above, this argument was inspired in part by my reading of Biggs' defence of the epistemic significance of phenomenal simulation (Biggs, 2009: pp. 654-657).

The Necessity of Acquaintance Claim: One cannot know what another person's experience of a mental state is like unless one has experienced a mental state of that sort for one's self.

A well-known application of this view is the Mary case that I discussed above. On the standard interpretation of that case, Mary (prior to her release) cannot know what it is like for other people to see a colour because she herself has never seen a colour. We could also apply this view to cases in which a person seeks to know what another person's affective states are like. For example, Kahane discusses the example of Zeno, a person who has never felt pain and so cannot know what it is like for others to feel pain (Kahane, 2010). A more everyday example is provided by the thought if you have not experienced depression (or anything sufficiently similar), you cannot know what it is like for another person to experience depression.

The second premise of the argument is that the best explanation of the Necessity of Acquaintance Claim is (what I call) the *Phenomenal Concepts Model*. This is a model of how we acquire phenomenal knowledge. It says that we acquire such knowledge through the use of phenomenal concepts, and that we can only acquire such knowledge through the use of phenomenal concepts.

A phenomenal concept is a concept that, in order to be deployed, requires the instantiation of the relevant phenomenal property (Loar, 1997; Papineau, 2002; Chalmers, 2003). For example, to deploy the phenomenal concept of red, I must instantiate redness by accessing a present moment experience of redness. I could do this by perceiving something red, by vividly remembering a red experience, or by imagining a red object. Similarly, to deploy a phenomenal

concept of sadness, I must instantiate sadness by accessing a present moment experience of sadness. Such an experience might be there already (if I am having a bad day), or I might generate it by accessing my memory, or by using one of the matching mechanisms associated with empathy.

The Phenomenal Concepts Model says that we can acquire phenomenal knowledge of the experiences of others by using phenomenal concepts. More specifically, it says that you can know what another person's experience is like just if you deploy the relevant phenomenal concept and combine it with a justified thought that the other person is experiencing something 'like this'. For example, if I want to know what it is like for a friend to see a bright blue sky, I can look at the same sky, attend to the colour that I am seeing, and think, 'my friend is experiencing something just like this', where the 'this' refers to the colour that I am perceiving. Similarly, if I want to know what it is like for a friend to experience depression, I can deploy a phenomenal concept by accessing a similar experience of depression in myself, and think 'my friend is experiencing something just like this'.

Moreover, the Phenomenal Concepts Model says that we can only acquire phenomenal knowledge in this way. Accordingly, it says that we cannot acquire phenomenal knowledge by using physical-functionalist concepts, which are the main other type of concept that we utilize. Physical-functional concepts are concepts that we deploy continually in everyday life to make sense of the world around us. To be deployed, they do not require the instantiation of any phenomenal properties. We may not experience anything when we deploy physical-functional concepts, or may we may just experience ourselves thinking of the associated word. If we use such concepts to understand another person's

mental state, we will frame that mental state either in physical terms (e.g. as a physiological or neurological state) or in functional terms (i.e. in terms of its likely causes and effects).

Prior to her release, Mary is an expert in using physical-functional concepts in order to understand the colour perceptions of other people. She can describe exactly what is going on in their eyes and in their brains, and can also tell you how colour perceptions are caused, and what their typical effects are. No matter how good she is at using these concepts, however, using these concepts will not give her knowledge of what it is like for other people to see colour. This is, according to the Phenomenal Concepts Model, because one cannot acquire phenomenal knowledge through using physical-functionalist concepts. Similarly, if a person was an expert in the psychology and neuroscience of depression, then this person could expertly deploy physical-functionalist concepts to understand the depression of other people. According to the Phenomenal Concepts Model, however, using such concepts will not allow this person to gain knowledge of what other people's experience of depression are like.

The claim that the use of physical-functional concepts cannot give rise to phenomenal knowledge is closely related to the claim that phenomenal properties are ineffable. To say that phenomenal properties are ineffable is to say that one cannot describe what they are like. For instance, one cannot describe what the colour red is like. If one could describe what the colour red is like, one could give Mary knowledge of what it was like before she had ever seen a colour. Similarly, one cannot describe what depression is like to someone who has never experienced it (or anything like it). To know what depression is like, one needs to

use the phenomenal concept of depression, which requires that one has access to an experience of depression.⁸⁶

The Phenomenal Concept Model explains the Necessity of Acquaintance Claim because it says that in order to know what another person's experience of a mental state is like, one must deploy the relevant phenomenal concept, and that to deploy the relevant phenomenal concept, one must be acquainted with the mental state in question. To illustrate, the Phenomenal Concept Model says that if I want to know what a friend's experience of seeing a blue sky is like, I must deploy a phenomenal concept of blueness, which I can only do by accessing a present moment experience of blueness. If I am accessing a present moment experience of blueness, however, I must be acquainted with such an experience.

Generally speaking, having had a prior experience of the relevant sort facilitates the use of phenomenal concepts. To continue the above example, if I have seen blue things in the past, it will be easier for me to access a present moment experience of blueness (through memory or imagination) than if I have never seen a blue object. Strictly speaking, however, such prior experience is not necessary for the deployment of a phenomenal concept. Even if I had never seen a blue object before, it is possible (though unlikely) that I might be able to accurately imagine seeing such an object. If I could do this, then I could deploy the associated phenomenal concept. Even in this case, however, I can still be

⁸⁶ To be more specific, I should say that an ineffable property is a property that one cannot describe directly, without using comparisons, metaphors or analogies. For instance, we might try to describe depression with a sentence that relates it to another feeling, such as: 'feeling depressed is like carrying around a ton of bricks'. However, we can only use such descriptions to know what depression is like if we are able to access an experience of the 'other feeling'. In other words, these descriptions work by helping us deploy the relevant phenomenal concept. For a discussion of the use of figurative language in understanding the pain of others, see Bending (2006).

said to be acquainted with the phenomenal property of blueness because I am experiencing it in the present moment. Similarly, if a person had never suffered from depression but was still somehow able to vividly and accurately imagine what depression felt like, then this person could deploy the phenomenal concepts associated with depression. We can say that they are acquainted with depression in the sense that they are accurately and vividly imagining being depressed.

So far, I have argued that the Phenomenal Concepts Model is the best explanation of the Necessity of Acquaintance Claim. The implication is that we should adopt the Phenomenal Concepts Model. The final step in my argument is to show that the Phenomenal Concepts Model entails the Strong Phenomenal Epistemic Claim that I am defending in this chapter.

To see this entailment, note first that even though the Phenomenal Concepts Model supports the Necessity of Acquaintance Claim, it contradicts the Sufficiency of Acquaintance Claim. It does this by saying that there can be cases in which one knows which affective state another person is experiencing, and in which one is acquainted with that affective state, but in which one fails to know what the other person's experience is like. Put simply, such a case can arise because one might fail to deploy the relevant phenomenal concept. In such case, even though one could deploy the relevant phenomenal concept (because one is acquainted with the associated experience), one does not.

To illustrate this point, consider an elaboration of the Mary case. Suppose that when Mary is first released, she is very taken with her new experiences of colour perception. Accordingly, she is often excitedly looking at new things and, when she is not, she is always imagining and remembering the colours that she

has seen. In this stage of her release, it is natural for her to access a relevant present moment experience of a colour whenever she is considering the colour experience of another person. She is, in short, always deploying phenomenal concepts and thus she is always acquiring phenomenal knowledge. For example, if a friend tells Mary about witnessing a purple sunset, Mary will eagerly comprehend the friend's experience by imagining or remembering a similar experience.

Years later, however, suppose that the novelty has worn off and that Mary has returned to the comfort of her black and white room, such that it has been a long time since Mary perceived a coloured object. Now, when she hears a friend describing a colour perception, she has reverted to understanding their mental state in the way that she used to, prior to her release. That is, she has gone back to relying on her physical-functionalist concepts. So when, for instance, her friend John tells her about how wonderful the blue sky looked this morning, Mary's mind expertly frames John's mental states in physical or functional terms. Mary does not, however, remember or imagine an experience of blueness. Accordingly, she cannot deploy the relevant phenomenal concept, and thus she lacks phenomenal knowledge of John's mental state. Even though Mary could, by utilizing her past experiences, form such phenomenal knowledge, she does not.

A similar story could be told about an expert in the psychology of depression who had never themselves experienced depression. In this experientially innocent state, this expert could only understand the depression of other people in physical-functional terms. If this person came to have their own experience of depression, then they might initially be struck with a sense of revelation about what other depressed people experience. The expert could use

their experience of depression to deploy the phenomenal concepts required to know what other people's experiences of depression feel like.

If, however, the expert stops feeling depressed, then they might revert to understanding other people's depression only through the use of physical-functional concepts. In this case, they would never use memory or imagination (or, affective contagion or perspective-taking) to access a present moment experience of depression, and so could never deploy the phenomenal concepts necessary to know what other people's experience of depression feel like. Even if the expert could deploy such concepts, they would not.

Indeed, there are two reasons that this kind of reversion to physical-functional concepts is more likely with respect to affective states than it is with respect to states of colour perception.⁸⁷ First, it is easier to generate novel experiences of colour perception than it is to generate novel affective states. If one wants to remind oneself what a blue sky looks like one can (assuming one is not confined to a black and white room) simply look outside on a clear day. On the other hand, it would be difficult (and very painful) to deliberately put oneself through another bout of depression in order to know what it feels like. Secondly, it is easier to recall or imagine experiences of colour perception than it is to recall or imagine experiences of painful affective states. People who have been through painful affective states generally do not want to remember what they were like.

⁸⁷ On the other hand, there is also a reason that it might be easier to deploy the phenomenal concepts associated with affective states than it is to deploy the phenomenal concepts associated with colour perception: the matching mechanisms of affective contagion and perspective-taking can help us generate the affective states required for deploying the phenomenal concepts associated with affective states.

In addition to contradicting the Sufficiency of Acquaintance Claim, the Phenomenal Concepts Model also directly supports the Strong Phenomenal Epistemic Claim. To see this, recall that the Phenomenal Concepts Model says that you will have phenomenal knowledge of another person's mental state just if you deploy the relevant phenomenal concept and combine it with a justified thought that the other person is experiencing something 'like this'.⁸⁸ On my account of empathy, deploying a phenomenal concept in this way with respect to another person's affective state is the same thing as empathizing with them. To empathize with another person's affective state is to use the relevant phenomenal concept to understand it. If the use of phenomenal concepts is the only way to acquire phenomenal knowledge of the affective states of others, then it follows that empathy is the only way to acquire phenomenal knowledge of the affective states of others.

In this section, I have offered an argument for the Strong Phenomenal Epistemic Claim. In short, that argument runs as follows: we should accept the Phenomenal Concepts Model because it is the best explanation of the Necessity of Acquaintance Claim, and the Phenomenal Concepts Model entails the Strong Phenomenal Epistemic Claim.

6. Two Possible Objections

So far in this chapter, I have been offering a defence of the Strong Phenomenal Epistemic Claim, which says that empathy is the only way through

⁸⁸ For a brief discussion of what justification here would, and would not, involve, see my discussion of the Phenomenal Attribution Condition in section 3 above.

which we can acquire phenomenal knowledge about the affective states of other people. I began that defence by defining empathy as the experiential understanding of the affective states of other people. I then gave two kinds of example which supported the idea that such understanding was the only way through which we could acquire phenomenal knowledge about the affective states of other people. Finally, in the last section, I offered an argument for the Strong Phenomenal Epistemic Claim.

In this section, I respond to two possible objections that might be made to my defence of the Strong Phenomenal Epistemic Claim. The first objection says that the Sufficiency of Acquaintance view remains appealing, and that it contradicts the Strong Phenomenal Epistemic Claim. The second objection says that we should not define empathy as experiential understanding. I now present, and respond to, each objection in turn.

6.1 The Sufficiency of Acquaintance Claim revisited

Earlier, I considered the *Sufficiency of Acquaintance Claim* about how phenomenal knowledge is acquired.

The Sufficiency of Acquaintance Claim: If a person B is feeling affective state S, then if person A has had at least one experience of an affective state that is similar to S, and if A knows that B is feeling S, then A knows what it is like for B to feel S.

I said that this claim was inconsistent with the Strong Phenomenal Epistemic Claim. A proponent of the Sufficiency of Acquaintance Claim may contend that I have not done enough to undermine its intuitive appeal, and thus maintain that we

should accept that claim, and reject the Strong Phenomenal Epistemic Claim. In response to this objection, I accept that there is something intuitively compelling about the Sufficiency of Acquaintance Claim, but argue that when that claim is formulated in its most plausible version, it is consistent with the Strong Phenomenal Epistemic Claim.

My suggestion is that these two claims are about a different kind of knowledge.⁸⁹ On the one hand, I suggest that the Strong Phenomenal Epistemic Claim should be interpreted as saying that empathy is the only route through which we can acquire *occurrent knowledge* of what another person's affective experience is like. Occurrent knowledge is knowledge that we are consciously attending to. On the other hand, I suggest that the *Sufficiency of Acquaintance Claim* should be interpreted as saying that acquaintance is sufficient for *dispositional knowledge* of what another person's affective experience is like. We have dispositional knowledge just when we have the disposition to form the relevant piece of occurrent knowledge when appropriately prompted.

If I am completing a physics exam, and I am given the question 'Who invented the Theory of Relativity?', I will (hopefully) form the occurrent knowledge that Albert Einstein did so. I already had the corresponding dispositional knowledge, and it was in virtue of having that dispositional knowledge that I was able to form the occurrent knowledge. If I was very stressed in the exam, I might fail to form the occurrent knowledge even though I had the dispositional knowledge, because my stress prevented the disposition from activating.

⁸⁹ My ideas in this section have been influenced by reading Tye (2012).

When it comes to having phenomenal knowledge of the affective states of other people, I shall assume that what matters is occurrent knowledge. If we are seeking to use phenomenal knowledge as an input into moral deliberation or as a way of offering understanding to a friend in need, I suggest that we need knowledge that we are consciously attending to. For simplicity, I shall continue to refer to this occurrent phenomenal knowledge simply as 'phenomenal knowledge'. It has been my claim in this chapter that empathy is the only route through which we can form such knowledge.

The intuition that the Sufficiency of Acquaintance Claim captures is that having one experience of an affective state typically puts us in a position to be able to access a similar experience again in the future (through memory or imagination). Consequently, it gives us the ability to form occurrent phenomenal knowledge of what other people are going through when they experience similar affective states. In other words, it gives us dispositional phenomenal knowledge. For example, if you have experienced depression in the past, you may well be in a position to remember or imagine those feelings in order to help you know what a depressed friend is going through. Whether you do so or not is a different question.

I say only that acquaintance typically gives rise to dispositional knowledge because there may be cases in which a person has experienced an affective state in the past, but cannot now remember or imagine such a state. This could be because the affective state was too long ago, or because it was too painful. By way of illustration, a survivor of intense torture may be unable to access any feelings similar to the intense physical and emotional pain that he went through.

In this case, he would be acquainted with those feelings, but he would not have dispositional knowledge of what they were like.

6.2 Challenging the definition of empathy

In order to defend the Strong Phenomenal Epistemic Claim, I defined empathy to be the experiential understanding of the affective estates of other people. Consequently, one might object to my approach by contending that empathy should be defined in a different way. On alternative definitions of empathy, the Strong Phenomenal Epistemic Claim may be false.⁹⁰

In response to this line of objection, I concede that I cannot offer conclusive reasons for preferring my definition of empathy to any other. People have used the term empathy in different ways, and I suspect that they will continue to do so. If one wishes to define empathy in a way that is different to the way I have defined it here, then one can instead take this chapter to be about the epistemic significance of experiential understanding, rather than being about empathy. That said, I do wish to briefly suggest one consideration in favour of my definition.

My impression is that, at least for most people, the idea of empathy is very closely connected with the idea of taking up another person's perspective. Frequently, philosophers have treated perspective-taking as a mechanism

⁹⁰ For example, if empathy is defined to be what I am calling 'sympathy', then it can occur without one knowing anything about the mental state of the other person (e.g. I could sympathize with a happy friend whom I falsely believe to be upset). Alternatively, if empathy is defined to be an affective match arising through emotional contagion, then I could empathize with another person without even being aware that I am doing so and without having any awareness of my empathic affective state.

through which an affective match can arise (see, e.g., Darwall, 1998; Coplan, 2011). The idea is that, through imagining being in someone else's situation, we can come to feel as they feel. If the notion of empathy is tied to this mechanism, however, then one will exclude the possibility that I could empathize with another person through affective contagion, or through cases (discussed above) where I match what they are feeling by drawing on my own life experiences.

On my account, this kind of perspective-taking has importance as a route to empathy, but empathy itself is to be identified with a different kind of perspective-taking. On my account, empathy consists in perspective-taking in the sense that it involves using your own affective state to understand the perspective of another person (that is, it involves forming phenomenal knowledge of what they are experiencing). So long as one's affective state is similar to that felt by the other person, it does not matter, on this approach, how one's affective state came about. It might have come about through the mechanism of perspective-taking, but it might also have arisen through affective contagion, or through one's own life experience. On this account of what perspective-taking involves, perspective-taking is an activity that you do with an affective state (using it to understand the perspective of another person), rather than an activity that gives rise to an affective state.

My suggestion is that defining empathy to be this kind of perspective-taking fits well with our everyday intuitions about what empathy is. Furthermore, as I have been arguing here, this kind of perspective-taking has a distinctive epistemic importance. As Smith (2017) notes, it makes sense to have a term that captures this phenomenon. If, on the other hand, one limited the term 'empathy' to cover only cases of affective matching that arose through the mechanism of

perspective-taking, then the term would fail to cover cases of experiential understanding (and thus of phenomenal knowledge) that arose through emotional contagion, or through the use of everyday life experiences.

7. Conclusion

This chapter has argued for the Strong Phenomenal Epistemic Claim which says that empathy is the only route through which we can acquire phenomenal knowledge about the affective states of other people. In making this argument, I offered a novel definition of empathy as experiential understanding. I also explored important (and overlooked) connections between empathy and the Knowledge Argument Debate.

In this chapter, I have made an important step forwards in the overall project of this thesis. That project is to show that empathy makes a distinctive contribution to moral deliberation. In the introduction of this thesis, I suggested that existing accounts of the role of empathy in moral deliberation were limited because they framed empathy only as a source of attributive knowledge. Such accounts are vulnerable to the objections that empathy is emotionally costly, and that it is source of normative bias. If empathy has these downsides, one could argue, one should instead prioritize alternative sources of attributive knowledge such as testimony and inference.

In this chapter, I have argued that empathy gives us more than attributive knowledge. To be more specific, I have argued that empathy, and empathy alone, gives us phenomenal knowledge of the affective states of other people. In the chapters to come, I argue that such phenomenal knowledge is a crucial input into

moral deliberation. In the next chapter, I argue that when we know what the suffering of another person is like, we will know how intrinsically bad their suffering is for them. Following that, in chapter 5, I argue that when we know what the suffering of another person is like, we will know the strength of the reason that we have to alleviate their suffering. Rather than being a source of normative bias, I shall argue, empathy is a source of normative insight.

Chapter 4: Empathy and Well-Being

1. Overview

The overall aim of this thesis is to demonstrate that empathy has an important role to play in moral deliberation. I began by arguing that the best explanation of empathy-induced altruism is the Phenomenal Epistemic Explanation, which says that empathy with a suffering person increases our altruistic motivation to help that person by giving us knowledge of what their suffering is like. A key component of the Phenomenal Epistemic Explanation is the claim that empathy is the only way in which we can have phenomenal knowledge about the affective states of other people. I have been calling this the Strong Phenomenal Epistemic Claim and in the previous chapter, I defended it. I now turn to the task of showing that phenomenal knowledge plays an important role in moral deliberation.⁹¹

In this chapter, I argue that, through teaching us what another person's suffering is like, empathy with that person also teaches us how bad that person's suffering is. To be more specific, I argue that empathy with a suffering person teaches us how bad, in itself, that person's suffering is for them. In other words, I shall say that empathy teaches us about the intrinsic prudential disvalue of the other person's suffering. I call this conclusion the Prudential Epistemic Claim. In section 2, I give a more detailed exposition of this claim.

I take it that people who are suffering are in either physical pain or emotional pain, or both. It is widely assumed that pain is intrinsically bad for the

⁹¹ For a different way of arguing that phenomenal knowledge is important to moral deliberation, see Hare (1981: ch. 5). I shall not discuss Hare's approach here because it is bound up with his meta-ethical view, which I do not share, that moral judgements are necessarily universalizable.

person who is experiencing it. My claim is that empathy is a way of finding out how bad another person's pain is. To contextualize and inform my argument for this claim, I will, in section 3, review the recent philosophical literature which discusses the badness of pain. In particular, I will look at what philosophers have had to say about how pain relates to suffering, about what makes pain bad, and about how we know that pain is bad.

In section 4, using insights from the aforementioned discussion, I offer an argument for the Prudential Epistemic Claim. I call this argument the *Argument from Introspection* because it relies on the idea that through introspecting our own suffering, we learn how bad it is (in itself, for us). Given that empathy involves such introspection, I will go on to argue, empathy inherits this epistemic power.

2. The Prudential Epistemic Claim

As I stated it above, the Prudential Epistemic Claim was focused on what we learn when we empathize with a person who is suffering. In this section, I shall start by discussing a more general version of the Prudential Epistemic Claim, which covers both cases of empathizing with a suffering person and cases of empathizing with a person who is experiencing positive affective states. I will then narrow down my focus to look specifically at cases of empathizing with a suffering person.

The general version of the Prudential Epistemic Claim says that when we empathize with another person with respect to an affective state that they are feeling, our empathy for them will give us knowledge of how good or bad their feeling is, in itself, for them. Where I say that a feeling is good or bad in itself, I

mean that it has *intrinsic* value. An affective state might also have *instrumental* value in so far as it has consequences that are intrinsically good or bad. For example, fear might feel bad, but have positive consequences in so far as it helps us to avoid dangerous situations. Back pain might be both bad intrinsically (it feels bad) and instrumentally (it can stop a person from sleeping). I am not claiming that empathy gives us knowledge of the instrumental value of the affective states of other people. Instead, my claim is that empathy gives us knowledge of the intrinsic value of the affective states of other people. In other words, empathy gives us knowledge of the value that those states have independently of their consequences.

Where a feeling is good or bad for the person who is experiencing it, I shall say that it has *prudential value*. A feeling that is good for a person has positive prudential value. By this I mean that it raises that person's well-being. In other words, we can say that this person's life is going better because they are having that feeling. Examples of good feelings include: excitement, gratitude, wonder, sexual pleasure, relaxation, and inspiration. If a feeling is good for me (if it has positive prudential value), then I have self-interested reasons to prolong it and to seek it in the future.

A feeling that is bad for a person has negative prudential value. By this I mean that it lowers that person's well-being. In other words, we can say that this person's life is going worse because they are having that feeling. Examples of bad feelings include: grief, fear, guilt, shame, headaches, and nausea. Having intense and persistent bad feelings can seriously detract from the quality of one's life. If a feeling is bad for me (if it has negative prudential value), then I have self-interested reasons to eradicate it and to avoid it in the future.

In addition to their prudential value, affective states can also have *moral value*. To see how moral value and prudential value might come apart, consider the example of a serial killer who has been placed in solitary confinement. If this person has feelings of loneliness, it seems clear that those feelings are prudentially bad (that is, that they make this person's life worse), but it could be debated whether these feelings are also morally bad, as some may regard them as being deserved.

In this chapter, I am arguing for the Prudential Epistemic Claim. The *General Version* of that claim can be stated as follows.

The Prudential Epistemic Claim (General Version): When a person A empathizes with a person B who is feeling an affective state S, then A's empathy for B will give A knowledge (if A did not have it already) of the intrinsic prudential value of S.

For example, if I empathize with the feelings of relief and excitement that a student has upon winning a place on a new course, I will know how good those feelings are, in themselves, for the student. Alternatively, if I empathize with the dismay of a student who has failed to win a place on a new course, I will know how bad those feelings are, in themselves, for the student. It should be emphasized that the Prudential Epistemic Claim says that empathy is sufficient for the knowledge in question, but not that it is necessary. This claim allows that there may be other routes to achieving such knowledge.

One possibility here is that another person could use introspection to discover how good or bad their own affective state is (see the Introspection Claim defended in section 4 below), and then tell me what they learn. If I could

understand and trust their testimony, I could thus come to know the intrinsic prudential value of their affective state without empathizing with them. Another possibility is that (perhaps through the use of introspection and empathy) a list could be developed which assigns an intrinsic prudential value to each affective state that a human can feel. If I had access to such a list, then if I knew which affective state another person was feeling, I could infer the intrinsic prudential value of that affective state without needing to empathize with the other person.

To clarify, the Strong Phenomenal Epistemic Claim (which I defended in the previous chapter) entails that if I do not empathize with the other person, I cannot know what their affective state feels like. However, the scenarios discussed in the above paragraph suggest that I could know the intrinsic prudential value of the other person's affective state without knowing what it was like.

As discussed earlier in this thesis, one might come to know which affective state another person is feeling without empathizing with them: one might listen to their testimony, or one might infer what they are feeling from what one knows about their behaviour or circumstances. To demonstrate the significance of the Prudential Epistemic Claim, I now describe several ways in which one might, through testimony or inference, know which affective state another person is feeling, but fail to know how bad that affective state is for them. The implication is that, according to the Prudential Epistemic Claim, empathy is special: unlike testimony and inference, empathy with another person guarantees that we will know the intrinsic prudential value of their affective state.⁹²

⁹² When I speak of testimony and inference in this sentence, I am referring to the use of testimony and inference to find out which affective state another person is feeling. As I said above, it might

First, there are cases in which one knows which affective states another person is feeling, but in which one fails to consider how those affective states impact that person's well-being. Such a case might arise, for example, if I am focused on how another person's affective states impact their ability to complete work that I need them to do. To illustrate, if a work colleague reports being depressed, then it is possible for me to view their depression solely through the lens of how it impacts me. In this case, thinking that I will now have more work to do, I may feel frustrated and stressed. The thought that their depression is bad for them might not occur to me because I am instead focused on how (instrumentally) bad their depression is for me.

Secondly, there are cases in which one considers how a person's affective state might have an impact on that person's well-being, but in which one is uncertain about what that impact is. In other words, there are cases in which one does not know the intrinsic prudential value of another person's affective state. Such a case might arise if one is trying to understand a person who is experiencing an affective state that one is not acquainted with (or that one has not experienced for a long time). By way of illustration: if one had never experienced intense anxiety, one might wonder how bad it feels. Similarly, if one is not prone to feeling anger, one might wonder whether its prudential value is negative or positive.

Finally, there are cases in which one forms a false view of the intrinsic prudential value of another person's affective state. Such a case might arise if

also be possible to use testimony and inference to find out the intrinsic prudential value of another person's affective state. The point that I am making here is that you can use testimony or inference to acquire the first piece of knowledge (about which affective state another person is feeling) without using it to acquire the second piece of knowledge (about the intrinsic prudential value of that affective state).

one adopted a false theory of well-being which assigned incorrect intrinsic prudential values to affective states. In an extreme case, one might adopt a theory of well-being that said that affective states have no intrinsic prudential value. Instead, such a theory might say that what matter's for a person's well-being is only that they succeed in achieving the goals that are important to them. If one held such a view, then one could know that another person was feeling anxiety, but fail to know how intrinsically bad that anxiety was for them because one would falsely believe that the anxiety was not intrinsically bad for them.⁹³

All of these cases illustrate the epistemic significance that the Prudential Epistemic Claim gives to empathy. According to that claim, empathy not only tells us which affective state another person is feeling; it also gives us the further knowledge of how intrinsically good or bad that feeling is for them. If, by contrast, we come to know what another person is feeling through testimony or inference (without thereby empathizing with them), then there is no guarantee that we will have this further knowledge.

Although we can also empathize with people who are feeling positive affective states, in this thesis I am focusing on cases of empathizing with those who are suffering. I shall soon say more about what suffering is. For now, an approximate definition will suffice: to suffer is to feel a painful affective state. Suffering can be physical (such as a headache, broken bone, or back pain) and it can be emotional (such as grief, shame and loneliness). Suffering is, I claim,

⁹³ The kind of theory of well-being that I am discussing here could still stay that affective states have instrumental prudential value, in so far as they helped one to achieve important goals. Accordingly, someone endorsing such a theory could believe that another person's anxiety was instrumentally bad for them, in so far as it prevented them from achieving their goals.

intrinsically bad for the person who is experiencing it. From now on, I will utilize the *Suffering Version* of the Prudential Epistemic Claim.

The Prudential Epistemic Claim (Suffering Version): When a person A empathizes with a suffering person B who is feeling S (an incidence of B's suffering), then A's empathy for B will give A knowledge (if A did not have it already) of the intrinsic prudential badness of S.

In other words, this claim says that when you empathize with a suffering person, you will know how bad their suffering is, in itself, for them.

For the purposes of this chapter, I shall assume that knowing the intrinsic prudential badness of the suffering of other people is important for moral deliberation. The basic idea behind this assumption is that our reason to help other people is determined (at least in part) by how badly off they are.⁹⁴ One way that a person can be badly off is by suffering (that is, by having painful affective states). Consequently, by telling us how badly off people are in virtue of their suffering, empathy helps us to recognize the strength of the reasons that we have to help those people. In the next chapter, I shall give a more extensive exploration of how empathy gives us knowledge about our reasons to help other people.

In section 4 below, I provide an argument for the Prudential Epistemic Claim. Before that, I lay important foundations for that argument by reviewing the contemporary philosophical discussion about the badness of pain. This will also allow me to contextualize my claim that suffering is bad, and to say more about what suffering is.

⁹⁴ In other words, this idea says that our reason to help other people is determined (at least in part) by how low their well-being is. It says that (as a general rule) if a person's wellbeing decreases, our reason to help them becomes stronger.

3. The Contemporary Philosophy of Pain

3.1 *The Badness of Pain*

Contemporary philosophers have tended to hold the view that pain has negative intrinsic prudential value.⁹⁵ Even if pain can have instrumental benefits, the consensus seems to be that it is, in itself, bad for us.⁹⁶ For an example of this view, consider the following account of pain, offered by Kahane.

It makes the person's life worse, a little bit worse if the pain is weak and brief, a lot worse if it's awful and long. Long periods of intense pain can really blight a life. And when the pain is intense and unremitting, utterly pervading the end of a life or even an entire life, some would say that the pain makes the life not just bad but not worth living. It reduces well-being to such an extent, it is thought, that even non-existence is better than existence. (Kahane, 2016: p. 209)

It is important to distinguish the claim that pain has negative intrinsic value from the more general doctrine of Hedonism. A *Hedonist Theory of Well-Being* states that only pleasure is intrinsically good for a person and that only pain is intrinsically bad for a person. This view has received extensive criticism.⁹⁷ However, one can believe that pain is intrinsically bad without endorsing such a view, because one can allow both that things other than pain can make a life worse, and that things other than pleasure can make a life better.

⁹⁵ As we shall see shortly, there are exceptions to this. If a philosopher has an Internalist definition of pain, but holds the Dislike View about why pain is bad, then in (unusual) cases where pain is not disliked, they will not judge it to be bad.

⁹⁶ One way in which pain can have instrumental value is by helping us to avoid harmful situations, as when the pain of burning compels us to remove our hand from a hot stove. It is also worth noting that one can think that pain is always prudentially bad whilst also thinking that pain can sometimes be morally good. Perhaps, for example, one might think that it is morally good for people to get the pain that they deserve.

⁹⁷ One influential objection to Hedonism is that it is a 'philosophy of swine' insofar as it cannot account for the importance of exercising our higher faculties. A second influential objection to Hedonism is Nozick's experience machine objection, which purports to show that, contrary to Hedonism, a pleasant life is not a good one if it is lived in simulated world. For more detail on these objections, and for responses to them, see Crisp (2006).

Although there has been widespread agreement that pain is bad, there has been a pervasive disagreement about what it is that makes pain bad. This disagreement has an important bearing on my defence of the Prudential Epistemic Claim. To explain this disagreement, I will begin by making some preliminary remarks about how philosophers have sought to define what pain is.

3.2 Defining Pain

First, it is helpful to note that, in their discussions of pain, philosophers have typically focused on writing about physical pains, and have given less attention to emotional pains. Although I take it that we can empathize both with the physical pains of another person, and with their emotional pains, I shall in this section replicate this focus on physical pains. Later on in this chapter, I will consider the extent to which what is said here about physical pains can also be said about emotional pains.

Focusing on physical pains allows us to take the question of ‘what counts as a pain?’ and frame it as ‘what makes a physical sensation count as a pain?’ *Internalists* answer this question by saying that it is the phenomenology of certain physical sensations that make them count as pains (Goldstein, 1989; Sumner, 1999; Rachels, 2000).⁹⁸ They are called Internalists because they take the defining feature of a pain to be one that is internal to it. Even if painful sensations

⁹⁸ Sumner introduces the Internalist-Externalist distinction, and also endorses the Internalist position. He traces this position back to the classical Utilitarians, and in particular to Bentham. For similar positions that have been adopted with regards to pleasures, see Kagan (1992), Crisp (2006) and Smuts (2011).

vary according to their bodily location, intensity, and type, the Internalist maintains that there is some phenomenal feature that they all share.

A different view is offered by *Externalists*.⁹⁹ An *Externalist* believes that what makes a sensation a pain is that we hold a particular kind of con-attitude towards it.¹⁰⁰ For simplicity of expression, I shall call this ‘particular kind of con-attitude’ *dislike*, although there is variation in how Externalists specify what the relevant kind of con-attitude is.¹⁰¹ It is because the dislike is external to the pain that the proponent of this definition is called an Externalist. One reason that philosophers are drawn to Externalism is that they are not convinced by the Internalist’s claim that there is some kind of phenomenological homogeneity amongst pains. Instead, the Externalist claims that pains are too diverse to be united in this way. What unites pains, for the Externalist, is that they are disliked.

3.3 The Phenomenal View

We can now turn to the question of what it is that makes pain bad.¹⁰² One view is that pain is bad because of what it feels like (Goldstein, 1989; Rachels, 2000).

⁹⁹ A third view, which I lack the space to discuss here, is the Body Damage View, on which pains are defined to be sensory representations of bodily damage (Nelkin, 1994; Tye, 1995: p. 113). I reject this view because I think that a pain remains a pain even when it is no longer associated with body damage. For example, consider the pain of a phantom limb. For further examples of pain without body damage, see Grahek (1991: p. 252 and p. 260). Even if one adopts the Body Damage View about the definition of pain, one can still accept the Hybrid View (defended below) about why pain is bad.

¹⁰⁰ Externalist views have often been developed in relation to pleasure, rather than to pain (Sidgwick 1981; Alston, 1968; Brandt, 1979; Carson, 2000; Heathwood, 2007). On an Externalist approach to pleasure, a pleasure is a sensation that we have a particular kind of pro-attitude towards (in simpler terms, it would be a sensation that we liked). I assume that philosophers who are Externalists about pleasure are also Externalists about pain.

¹⁰¹ For a useful overview of con-attitudes that we can have towards pain, see Rachels (2000: pp. 191-195).

¹⁰² Throughout this section, I take the term ‘bad’ to mean ‘intrinsically and prudentially bad’.

The Phenomenal View: A pain sensation is bad because of what it is like.

Those who hold the Phenomenal View also hold the Internalist definition of pain.¹⁰³ That is, they believe there is some phenomenological feature that all pains share, and they believe that pains are bad because of their phenomenology.

Defenders of the Phenomenal View claim that it is intuitively compelling (see, e.g., Goldstein, 1989: p. 255). They suggest that when we introspect pains, we draw the conclusion that they are bad, and that when we are prompted to explain this conclusion we typically say simply that the pain ‘feels bad’. Defenders of the Phenomenal View can acknowledge that we usually dislike pains. Indeed, they can both explain and justify this dislike: we dislike pains because they feel bad. On the Phenomenal View, the badness of pains is objective in the sense that it does not depend on the evaluative stance we take towards them: on the Phenomenal View a pain is bad whether we dislike it or not.

An important problem for the Phenomenal View is what I shall call *Exception Cases*.¹⁰⁴ An Exception Case is a case in which a person experiences a painful sensation but reports that they do not mind it. These cases are sometimes referred to as cases of ‘pain asymbolia’. To use an oft-quoted example, patients who have had a frontal lobotomy have described being in pain,

¹⁰³ Kahane labels the combination of these two views as ‘Sensation Theory’ (Kahane, 2016: p.210). If one holds the Phenomenal View, one has at least one promising option for defining pain in an Internalist way: one can say that a pain is any sensation that is bad because of what it is like.

¹⁰⁴ As Kahane (2016: pp. 210-211) notes, the Phenomenal View may previously have also been deemed to be unattractive because it stood in opposition to several philosophical approaches that have since declined in popularity. These approaches are: behaviourism (which involved scepticism about the phenomenal element of mental states), anti-objectivism about moral value, and the desire satisfaction account of well-being.

but not being bothered by it.¹⁰⁵ My experience has been that practicing some forms of meditation can have a similar (though less pronounced) effect.¹⁰⁶ Special circumstances may also lead a person to not mind their pain. For example, it has been suggested that some soldiers in WW2 field hospitals did not mind the pain of their injuries because they associated their injuries with the possibility of going home (Beecher, 1959). Masochists also perhaps provide examples of Exception Cases: masochists sometimes feel pain sensations and like them.

The Phenomenal View struggles to make sense of these cases. To see why, recall that the Phenomenal View is committed to saying that pains are bad because of their phenomenology, and thus that two pains with the same phenomenology must be equally bad. Consequently, if I have a pain which is bad and which has the same phenomenology as the pain experienced by (for example) a frontal lobotomy patient, the Phenomenal View implies that our pains must be equally bad and thus that the pain of the frontal lobotomy patient must also be bad. Yet, if we are to trust the self-report of frontal lobotomy patient, their pain is not bad.¹⁰⁷

¹⁰⁵ A frontal lobotomy involves severing many of the neural connections to and from the pre-frontal cortex of the brain. Now widely discredited, it was previously used as a treatment for some psychiatric illnesses in the 1940s and '50s. Grahek (2007) offers a useful survey of the scientific literature discussing how lobotomy patients relate to pain.

¹⁰⁶ During several 10 day vipassana meditation retreats, I was asked to perform regular sessions of sitting without moving for one hour and objectively observing the pain sensations that arose. As I progressed with this practice, I found myself being less reactive my pain, such that it seemed less bad to me. For a related account of meditation, see 'Chapter 9: Turning Towards Difficulties' in Williams and Penman (2011). Hilton et al (2017) offers a recent review of research into the use of meditation to manage chronic pain.

¹⁰⁷ A defender of the Phenomenal View could try to argue that the pain of the lobotomy patient is bad, by denying the accuracy of that patient's self-reports. This strategy, however, is problematic, because defenders of the Phenomenal View typically appeal to the reliability of introspection as a way of finding out about the value of our pains (see section 4 below).

3.4 *The Dislike View*

This line of reasoning has led many philosophers to adopt a different view about the badness of pain, on which pain is only bad when (and to the extent that) it is disliked (Brandt, 1979; Brink, 1997; Sumner, 1999).

The Dislike View: A pain sensation is bad because (and to the extent that) it is disliked.

This view can explain the badness of pain by appealing to a subjectivist theory of well-being: where we dislike a thing, our dislike of that thing makes that thing bad for us.

The Dislike View has no trouble accommodating Exception Cases. It says that the sensations in those cases are not bad because they are not disliked. Most philosophers who hold the Dislike View employ an Externalist definition of pain. On this definition, the sensations in the Exception Cases do not actually count as pains, because they are not disliked. Accordingly, someone who holds both the Dislike View and the Externalist definition of pain can maintain that pain is always bad. On this combination of views, even though the badness of pain depends on its being disliked, pain is necessarily disliked.¹⁰⁸ However, the Dislike View can also be combined with an Internalist definition of pain (Sumner, 1999). On this combination of views, pain is not always bad: the sensations in Exception Cases will still be deemed to be pains, but they will not be deemed to be bad, because they are not disliked.

¹⁰⁸ A question remains about whether the Externalist can maintain that pain is intrinsically bad. For a discussion of the corresponding question about whether an Externalist can maintain that pain is intrinsically good, see Feldman (1997).

3.5 Empathy and pain

The debate between the Phenomenal View and the Dislike View has an important bearing on my project in this chapter. If one accepts the Phenomenal View then, I suggest, the Prudential Epistemic Claim follows. To see this, recall that, by the argument of the previous chapter, if we empathize with another person's pain, we will know what it is like. If the Phenomenal View is true, then we can say that a pain's phenomenal feel (what it is like) determines its intrinsic prudential badness. It then seems a small step to say that if (through empathy) we know what another person's pain is like, we will know the intrinsic prudential badness of their pain.¹⁰⁹

If one accepts the Dislike View, however, my project is threatened. On the Dislike View, to know how bad a person's sensation of pain is, we must know whether or not they dislike it. Moreover, if they do dislike their pain, we must also know the extent to which they dislike it. Empathy does not give us this knowledge. Through empathy, we could know the phenomenology of the other person's pain without knowing whether (or how much) they dislike it. For instance, I could know exactly what a friend's back pain feels like, but fail to know his reaction to it and thus fail to know how bad it is for him. Consequently, if the Dislike View is correct, we could empathize with a person without knowing how bad their pain is, and thus the Prudential Epistemic Claim would be false. The worry here is not just that we might be empathizing with someone whose pain happens to be an Exception Case (e.g. someone who has had a pre-frontal lobotomy, or someone who has

¹⁰⁹ The argument given here is only an approximate version of the argument that I give for the Prudential Epistemic Claim in section 4 below. As stated here, the argument is invalid, because one can know fact A (what a pain is like) and know that fact A determines fact B (how bad the pain is) without thereby knowing fact B.

done decades of rigorous meditation practice). Rather, it is possible that ‘normal’ people all vary in the extent to which they dislike pains. Even if we can reliably guess that the other person dislikes their pain, we still face the further challenge of working out how much they dislike their pain.

One way of defending the Prudential Epistemic Claim would be to try to defend the Phenomenal View against the Dislike View. I shall not do that here. Instead, the strategy that I shall pursue in this chapter is to follow Kahane (2009, 2016) in adopting a third view about why pain is bad. This view combines the strengths of the Phenomenal View and the Dislike View, and thus I shall call it the *Hybrid View*.¹¹⁰

3.6 The Hybrid View

The first step to make in developing the Hybrid View is to switch from explaining the badness of pain to explaining the badness of pain-combined-with-dislike. I shall refer to this combination of pain and dislike as *suffering*.¹¹¹

The second step is to posit that the phenomenology of this composite state is a combination of the phenomenology of the physical sensation and of phenomenology of the dislike of that sensation. This step assumes that our dislike of pain is a consciously felt mental state that has its own phenomenology.¹¹² One

¹¹⁰ Kahane refers to it as the ‘Felt Aversion Theory’.

¹¹¹ In using the term ‘suffering’ in this way, I am following Hare (1972). For an Externalist, all instances of pain will be part of an instance of suffering, because they will necessarily be disliked. For an Internalist most pains will be part of an instance of suffering, but Exception Case pains will not be, because they are not disliked. Kahane refers to painful body sensations as ‘sensations of pain’ and refers to the combination of painful sensations and dislike as ‘unpleasant experiences’. As a shorthand, he refers to the latter simply as ‘pain’ (Kahane, 2016: p. 214).

¹¹² I do not mean to say that we cannot dislike things subconsciously. There may be different types of dislike, some of which are conscious and some of which are unconscious. The Hybrid View says that the type of dislike that plays a critical role in determining the badness of our pain is

reason to adopt this assumption is that it can explain the epistemology of suffering: if dislike is a consciously felt mental state, then we can observe it through introspection and thus we can know when we are experiencing suffering (that is, we can know when we are experiencing a pain that we dislike).

The third, and final, step in developing the Hybrid View is to assert that suffering is bad because of its phenomenology.

The Hybrid View: Suffering (the combination of a pain sensation and the consciously felt dislike of that sensation) is bad because of what it is like.

The phenomenal feel of suffering might be bad for two reasons. First, if the Phenomenal View is true, then pain sensations are bad because of how they feel. The implication of this is that suffering will feel bad because it involves pain sensations. Secondly, even if the Phenomenal View is false, suffering will still feel bad because it involves the dislike of a pain sensation, and such dislike feels bad. In support of this latter point (that dislike feels bad), Kahane cites evidence of people who report feeling dislike without feeling a discernible physical pain, and who nevertheless find their experience to be unpleasant (Kahane, 2009: p. 335). In particular, Kahane draws on Ploner et al. (1999) to offer an account of a patient who had suffered damage to the part of his brain associated with processing touch sensations.

Such a patient reported a ‘clearly unpleasant’ feeling located ‘somewhere between his fingertips and his shoulder’ that he wanted to avoid, but was completely unable to further describe its quality, localization or intensity.¹¹³

conscious dislike. For a discussion of unconscious dislike (and unconscious ‘liking’), see Berridge and Winkielman (2003).

¹¹³ In this passage, I am citing Kahane, who is in turn (within the quotation marks) citing Ploner.

Using the approach developed above, we can say that this patient was having a bad experience, and that his experience was bad because it involved feeling dislike towards a pain sensation.

The Hybrid View can allow that the badness of suffering increases as the dislike involved in suffering becomes stronger. In this way, it can allow that dislike determines the badness of suffering. It contends, however, that the way in which dislike does this is by changing the phenomenology of suffering. The phenomenology of suffering is determined by the type of pain sensation involved, and by the strength of dislike involved. In other words, the Hybrid View says that dislike makes suffering bad by making it feel bad. An implication of this approach is that if one is suffering, one can improve the quality of one's experience by finding ways to reduce one's aversive reactions to pain.

The Hybrid View retains the intuitive appeal of the Phenomenal View because it says that suffering is bad because of how it feels. It can explain and justify our opposition to suffering: it says that we have reason to minimize suffering in our lives because it feels bad. Given that we typically dislike our pain sensations, the Hybrid View can also rationalize our opposition to pain: it says that we have reason to avoid pain because when we feel pain we tend to suffer (and suffering feels bad).

The Hybrid View also incorporates the appeal of the Dislike View because it can accommodate the Exception Cases. The Hybrid View can allow that pain, when it is not disliked, is not bad.¹¹⁴ Moreover, the Hybrid View can be used to

¹¹⁴ In other words, the Hybrid View (which is about the badness of suffering and not about the badness of pain) is consistent with the rejection of the Phenomenal View. As I noted above, it is also consistent with the acceptance of the Phenomenal View.

explain why such pain is not bad: it is not bad because it does not involve the 'bad phenomenology' associated with dislike.

It is important to note that, on the Hybrid View, there are two ways in which we can be against pain. One way, described above, involves introspecting our experiences of pain and assessing their phenomenal qualities. It involves noticing that instances of pain are typically instances of suffering (because we typically dislike our pain) and that such instances of suffering are bad because of how they feel.

The second way in which we can be against pain is through our pre-reflective dislike of pain, which cannot itself be explained (or justified) in terms of the badness of pain. Indeed, if we take the Exception Cases seriously, we must acknowledge that pain itself (in isolation from dislike) is not bad. The implication is that our pre-reflective dislike of pain is in some sense a mistake, albeit one that might be instrumentally useful (our aversion to pain might help to protect us from bodily damage). Nevertheless, where such dislike arises (and it is, I guess, hard to eradicate), it combines with pain sensations to create experiences of suffering, which are bad and which are bad because of how they feel.

My aim in this section has been to summarize Kahane's argument that the Hybrid View is a superior alternative to the Dislike View. The Hybrid View offers an approach that can both retain the intuitive appeal of the Phenomenal View and accommodate the Exception Cases. In section 4, I will use the Hybrid View to construct my argument for the Phenomenal Epistemic Claim. Roughly speaking, the logic will run as follows: if the phenomenology of a person's suffering determines how bad it is, then where we use empathy to know the

phenomenology of another person's suffering, we will thereby come to know how bad their suffering is.

3.7 Empathy with suffering

Before giving my argument for the Prudential Epistemic Claim in more detail, it is worth pausing to reflect on the implications that the preceding discussion of dislike has for empathy. I have said that the phenomenology of a person's suffering is determined both by the phenomenology of their pain and by the phenomenology of their dislike of their pain. Consequently, if one wants to know what another person's suffering is like by empathizing with them, it will be necessary not only to match their pain, but also to match the level of dislike that they have towards their pain. By way of illustration, if you try to replicate the experience of a pre-frontal lobotomy patient with back pain by recollecting an experience of back pain that you detested, you will fail. The reason that you will fail is that the experience of the other person will be very different from your experience, because you disliked your pain but they did not dislike their pain.

We can now see that it is important to distinguish between empathizing with another person's physical suffering and empathizing with their physical pain. The former involves matching both the other person's pain and the level of dislike that they have towards that pain. Such empathy will still give rise to knowledge of what the other person's experience is like. Empathising with another person's pain, however, merely involves matching their pain and does not give rise to knowledge of what the other person's experience is like.

Throughout this entire section, I have been focused on the analysis of physical pain and suffering. I adopted this focus from the philosophical literature that I was discussing. Whilst we can empathize with the physical suffering of other people, however, it is more common to talk about empathizing with the emotional suffering of other people. Accordingly, I am now faced with an important question: how does the above analysis transfer over to emotional suffering?

My suggestion is that the Hybrid View can also be applied to cases of emotional suffering. I take an incidence of 'emotional suffering' to be an emotional pain (such as depression, anxiety, self-criticism, and so on) combined with a consciously felt dislike of that state. Reflections on my own experiences of emotional pain suggest that I hold such dislike towards it. The 'emotional version' of the Hybrid View would then say that emotional suffering is bad because of how it feels. This might be because painful emotions feel bad in themselves, even if they are not disliked.¹¹⁵ Alternatively, it could be because it feels bad to dislike a painful emotional state. If dislike plays an important role in determining the phenomenology and the badness of emotional suffering, then it will be important to empathize with the emotional suffering of other people, rather than merely empathizing with their emotional pain.

Before concluding this section, it is important to note that matching the dislike of another person presents an additional challenge to the would-be empathizer. If we want to empathize with the suffering of other people, however,

¹¹⁵ This would involve endorsing an 'emotional version' of the Phenomenal View. Such a view may be easier to defend than the Phenomenal View about physical pain, because it does not (yet) face the challenge of Exception Cases in which people feel emotional pain but report not minding it. One might also be attracted to the Phenomenal View about emotional pain if one thinks that the dislike involved in physical suffering is a negative emotion (such as fear or resentment). In that case, saying that the dislike feels bad in itself would commit one to saying that negative emotions can, in themselves, feel bad.

this is a challenge that we must engage with. In fact, I now suggest, there are actually two challenges here.

First, there is the *Epistemological Challenge*: how do we find out about the level of dislike that the person we are empathizing with has towards their pain? We might be able to use our reactions to our own affective states as a guide to how others are likely to react: if I hate the pain of having a dental cavity, you probably do too. Yet it is possible that you react to it in a different way than I do. The worry here is not just that you might have had a pre-frontal lobotomy, but that you might have found a way to be more accepting of physical pain than I am. Alternatively, it is possible that you might be more reactive to dental pain than I am (e.g. because you had awful experiences with dentists as a child).

Secondly, there is the *Psychological Challenge*: even if I know the extent to which you dislike your pain, am I able to replicate your dislike? This challenge will be most problematic in cases where we are trying to empathize with someone who has a pattern of reactivity that is very different from our own. For example, consider the challenge I would have in empathizing with a person who has had a pre-frontal lobotomy. To empathize with the pain of such a person, I would need to feel a similar pain and to not mind my pain. This will be difficult, even impossible for me to achieve. Reversing the situation, it may also be difficult for the person who has had a pre-frontal lobotomy to empathize with the physical suffering others. This person could perhaps feel the pain of other people, but would struggle to replicate the dislike that other people have towards their pain.

With this preparatory discussion complete, I now move to offer my argument for the Phenomenal Epistemic Claim. I call it the *Argument from Introspection*.

4. The Argument From Introspection

4.1 Recap

To begin, it will be helpful for me to restate the claim that I am arguing for in this section.

The Prudential Epistemic Claim (Suffering Version): When a person A empathizes with a suffering person B who is feeling S (an incidence of B's suffering), then A's empathy for B will give A knowledge (if A did not have it already) of the intrinsic prudential badness of S.

As discussed above, I take suffering to be a composite state involving pain (either physical pain or emotional pain) and a consciously felt dislike of that pain.

It will also be helpful for me to recap what empathy with suffering involves. I shall do this by applying the definition of empathy that I offered in the previous chapter to the case of empathizing with another person's suffering. Taking empathy to be the experiential understanding of the affective states of other people, I said that there were three conditions that must be met for empathy to be taking place. The first condition is the Affective Match Condition.

The Affective Match Condition: Person A must feel S* where S* is similar to S.

In the previous section, I said that in order to match the suffering of another person, one must both feel a pain which is similar to the pain felt by the other person and one must also feel a level of dislike towards one's pain that is similar to the level of dislike that the other person has towards their pain.

To meet the second condition, a would-be empathizer must introspect their suffering.

The Introspection Condition: Person A must introspect S*.

In other words, person A must introspect their composite state of having a pain and disliking it. This introspection will give them phenomenal knowledge of what their suffering feels like.

To meet the third condition, a would-be empathizer must use their experience as a representation of what the other person is going through.

The Attribution Condition: Person A forms a conscious and justified belief that person B is experiencing something 'like this' where 'this' refers to the phenomenal quality of S*.

Having restated the Phenomenal Epistemic Claim and the account of empathy that I will use to defend it, I now proceed to offer the Argument from Introspection. That argument has two premises. I introduce each in turn.

4.2 *The Introspection Claim*

The first assumption is a claim about what we learn when we introspect our own suffering. Put simply, this claim says that through introspection we can learn how bad our suffering is.¹¹⁶

The Introspection Claim: Where a person A introspects an incidence of their own suffering, S^* , A's introspection of S^* will give A knowledge of the intrinsic prudential badness of S^* .

The implication of this claim is that when I empathize with another person's suffering, and thus feel a similar suffering (S^*), I will know the intrinsic prudential badness of my own suffering. For example, if I am stressed about an upcoming exam and I introspect my stress, the Introspection Claim says that I will know the intrinsic prudential badness of my stress.

I have come to accept the Introspection Claim through reflecting on what I learn when I introspect my suffering. Indeed, I take the claim to be self-evident to anyone who is introspecting their suffering. In other words, I suggest that when one is introspecting one's suffering it is self-evident that one knows, through one's introspection, how bad one's suffering is.¹¹⁷

It is important to make several clarifications about the Introspection Claim so that it may be protected against obvious counter-examples. First, when considering composite experiences such as suffering, the Introspection Claim

¹¹⁶ For endorsements of the idea that introspection can tell us about the badness of pain (or about the goodness of pleasure), see Goldstein (1989: p. 255), Mendola (1990), Sinhababu (2012: pp. 17-22) and Kahane (2016: pp. 217-8).

¹¹⁷ It should be noted that on the Dislike View of pain that I discuss above, the Introspection Claim may be false because introspecting an incidence of suffering may not give us knowledge of the extent to which we dislike the pain (Kahane, 2009: p. 335). This is because on the Dislike View, dislike need not be consciously felt. I take this is an additional reason to reject the Dislike View.

says that introspection will give us knowledge of how bad the overall experience is, but not that it will give us knowledge of which components of the experience make the overall experience bad. To illustrate this point, suppose that I am suffering from toothache. According to the Introspection Claim, if I introspect my suffering, I will know how bad my suffering is. The Introspection Claim does not say, however, that I will know which part of the suffering makes it bad. It would be consistent with the Introspection Claim for me to mistakenly believe that it is the pain sensation that makes the suffering feel bad, when it is in fact my dislike of the pain sensation that makes the suffering feel bad.

Secondly, the Introspection Claim does not entail that all self-reports of well-being are accurate. Rather, it implies only that a person's self-report would be accurate if they had introspected on their present affective state and honestly communicated their subsequent learning to you. If a person making a self-report about their well-being does not introspect their present feelings, or if they lie, then the Introspection Claim says nothing about the veracity of what they say. For example, a person who is feeling unusually happy might report that they are not doing well, because they are so used to that being the case (if they had introspected their happiness, they would have realized the upturn in their well-being). Alternatively, a depressed person may know full well (through introspection) that their well-being is low, but tell other people that they are feeling fine.

Thirdly, it should be emphasized that the Introspection Claim says only that introspection is a way of knowing about the value of our present-moment affective states. The Introspection Claim does not say anything about how good we are at assessing the value of an experience that takes place over an extended period of

time. Accordingly, it would be consistent with the Introspection Claim to say that we routinely make errors in such evaluations.¹¹⁸ The Introspection Claim also does not say anything about our ability to evaluate our past affective states.

Taken by itself, the Introspection Claim does not entail the Prudential Epistemic Claim. On the basis of the Introspection Claim, all we can say is that when I empathize with the suffering of another person, I will introspect an incidence of suffering that is similar to that felt by the other person, and that I will thus know how bad my suffering is. The Introspection Claim does not say anything about what I will know about the value of the other person's suffering. To establish the Prudential Epistemic Claim, I must answer the following question: how can the empathizer move from knowing the value of her own suffering to knowing the value of the other person's suffering?

4.3 The "Self-Evident Supervenience" Claim

The second premise in the Argument from Introspection is, like the first, about what we learn when we introspect our suffering. Building on the first premise (which says that we learn how bad our suffering is), this second premise says that when we introspect an incidence of our suffering we will know that it is bad because of how it feels (such that any experience of suffering that felt the same way would be equally bad). In more technical terms, this premise says that

¹¹⁸ For example, it has been shown that adding a better (though still bad) end to a bad experience makes it more preferable to subjects. In one experiment, subjects preferred (i) having a hand in painfully cold water for sixty seconds followed by having it in slightly less cold water for thirty seconds to (ii) having a hand in painfully cold water for sixty seconds (Kahneman et al, 1993).

when we introspect an incidence of our suffering we will know that its intrinsic prudential badness supervenes on its phenomenology.¹¹⁹

Because this second premise says that the supervenience relationship (between the phenomenology of suffering and its badness) is self-evident to anyone who introspects their suffering, I shall call it the “*Self-Evident Supervenience*” Claim.¹²⁰

The “Self-Evident Supervenience” Claim: When person A introspects on S* (an incidence of their suffering) and believes that someone else is feeling S (a similar incidence of suffering), A will know that S has the same intrinsic prudential badness as S*.

To illustrate this claim, let us return to the example above. According to the Introspection Claim, if I introspect my own stress as part of empathizing with a similarly stressed friend, I will know how bad my stress is. The “Self-Evident Supervenience” Claim adds that I will also know that my friend’s stress is just as bad as mine, because I will know that it feels just like mine.¹²¹ In other words, the “Self-Evidence Supervenience” Claim rules out the possibility that I will think that my friend’s stress has any more or less value than my own.¹²²

¹¹⁹ For an endorsement of the claim that the value of an experience supervenes on its phenomenology, see Kahane (2009: pp. 333-334).

¹²⁰ I do not mean to suggest that this supervenience relationship is self-evident to us even when we are not introspecting our suffering. A person engaged in abstract debates about this relationship might deny that the value of an incidence of suffering supervenes on its phenomenology. Indeed, as I note below, someone who holds the Dislike View might make such a denial. My point is that such a person could be convinced to correct their view by introspecting on an incidence of their suffering.

¹²¹ Sinhababu suggests a similar approach in relation to how we can know about the goodness of other people’s pleasures. He writes: “if in a pleasant moment I consider what it’s like for others to have exactly the experience I’m having, I must think that they’re having good experiences” (Sinhababu, 2012: pp. 22-23). It should be noted that Sinhababu here takes goodness to be ‘moral goodness’, as opposed to the ‘prudential goodness’ that I am talking about in this chapter.

¹²² The example given here is a simplification in so far as it does not discuss the dislike that my friend and I have towards our stress. As noted above, to empathize with my friend’s suffering, I will

The “Self-Evident Supervenience” Claim is supported by the Hybrid View that I outlined and defended above. That view says that suffering is bad because of how it feels. In other words, it says that the badness of suffering supervenes on its phenomenology. By contrast, a defender of the Dislike View may be sceptical of the “Self-Evident Supervenience” Claim. On the Dislike View, the badness of suffering is determined by the badness of the pain that it involves, which is in turn determined by the extent to which the pain is disliked. This dislike may or may not be consciously felt. Accordingly, increasing the extent to which a pain is disliked would make an incidence of suffering worse, but may not change its phenomenology. This possibility contradicts the idea that the value of an incidence of suffering supervenes on its phenomenology. Accordingly, the Dislike View is inconsistent with the “Self-Evident Supervenience” Claim, and it was for this reason that I argued against the Dislike View.

If my definition of empathy is combined with the two premises outlined above, the Prudential Epistemic Claim can be derived. According to my definition of empathy, when I empathize with the suffering of another person I will feel a similar suffering, I will introspect my suffering, and I will have a conscious and justified thought that the other person is feeling something ‘like this’. According to the Introspection Claim, it follows that I will know the intrinsic prudential badness of my suffering. According to the “Self-Evident Supervenience” Claim, it follows that I will know that the other person’s suffering is just as bad as mine. Thus,

need to match the dislike that he has towards his stress. I might be unable to empathize with my friend because I do not know how much he dislikes his stress. The point in this section is that if I am able to empathize with my friend, by recreating his suffering, then I will know that his suffering is just as bad as mine.

when I empathize with another person's suffering, I will know how intrinsically bad their suffering is for them.

5. Conclusion

This chapter has argued for the Prudential Epistemic Claim, which says that empathy with a suffering person gives us knowledge of how intrinsically bad their suffering is for them. After explaining that claim and contextualizing it against the recent philosophical discussions of the badness of pain, I defended it with the Argument from Introspection.

I introduced the Prudential Epistemic Claim as part of my project to show that empathy has an important role to play in moral deliberation. The thought was that by showing us how badly off a suffering person is, empathy with such a person helps us to learn about the strength of reason that we have to help them. Without empathy we might understate our reasons to help such a person, because we do not realize how bad their suffering is.

On this account of what we learn in empathy, however, there remains a gap between an empathizer knowing how bad the other person's suffering is, and the empathizer knowing the strength of their reason to help the other person. It would be possible for an empathizer to question the strength of the reason that they had to help the other person. They may even say, 'Yes, I know how bad this person's suffering is, but do I have any reason to help them?' In other words, if all empathy told us was how bad another person's suffering was, it may be possible for an empathizer to remain indifferent towards a suffering person.

In the next chapter, I argue that there is no such gap. I claim that empathy, in addition to giving us knowledge of how bad another person's suffering is, also gives us knowledge of how strong our reasons are to alleviate that person's suffering. Moreover, I will continue, empathy shows us that those reasons are stronger than we normally think.

Chapter 5: Empathy and Normative Reasons

1. Overview

In this thesis, my overarching aim is to show that empathy has an important role to play in moral deliberation. According to the standard Epistemic Defence of empathy (that I discussed in chapter 1), empathy contributes to moral deliberation by giving us attributive knowledge of what other people are feeling. This defence is vulnerable to the objection that testimony and inference serve as alternative ways of acquiring attributive knowledge that are superior to empathy in so far as they are less emotionally costly, and in so far as they avoid the tendency that empathy has to lead our decision-making astray. Consequently, I have been seeking to strengthen the Epistemic Defence of empathy by arguing that empathy contributes more than attributive knowledge to moral deliberation.

In chapters 2 and 3, I introduced, and then argued for, the claim that empathy (and empathy alone) gives us phenomenal knowledge about the affective states of other people. I called this the Strong Phenomenal Epistemic Claim. In addition to giving us knowledge of which affective states other people are feeling, this claim says, empathy also gives us knowledge of what it is like for other people to have such feelings. Moreover, the claim adds, empathy is the only way for us to acquire such knowledge.

Having defended the Strong Phenomenal Epistemic Claim, I was left with a question: in what way is phenomenal knowledge important to moral deliberation? In the previous chapter, I took a step towards answering that question by arguing that empathy for a suffering person gives us knowledge of how bad their suffering is, in itself, for them. I called this the Prudential Epistemic Claim. More generally, I

argued that empathy for another person is a way of finding out about the intrinsic significance of that person's affective states for their well-being.

In this chapter, I complete my defence of empathy by arguing that there is an additional piece of knowledge that empathy gives us. Roughly speaking, my claim is that empathy with a suffering person gives us knowledge of the strength of reason that we have to help that person. More specifically, I argue that empathy with a suffering person gives us knowledge of the strength of the *purely phenomenal reason* that we have to alleviate that person's suffering.

I shall begin, in section 2, by explaining the notion of a purely phenomenal reason. We can have many different types of reason to help a suffering person and a purely phenomenal reason is the reason to help them that is provided solely by the phenomenal quality of their suffering. When we help a person for a purely phenomenal reason, we can say that we helped them simply because they were experiencing something awful.

Using this terminology, I then (in section 3) formulate the two claims that I defend in this chapter. The first, the *Strong View*, says that we have purely phenomenal reasons to alleviate all instances of suffering, and that these reasons are stronger than common sense morality leads us to believe. The second says that we can learn about the strength of our purely phenomenal reasons by empathizing with people who are suffering. I call this view the *Normative Epistemic Claim* because it says that empathy gives us knowledge of our normative reasons for action. More specifically, the Normative Epistemic Claim says that when we empathize with a suffering person, our empathy for that person will give us knowledge of the strength of our purely phenomenal reason to

alleviate that person's suffering. If we combine the Strong View and the Normative Epistemic Claim, we can conclude that empathy will typically induce normative learning in the empathizer: through their empathy with a suffering person, the empathizer will learn that they had more reason to help the suffering person than they had previously thought.

In section 4, I offer a defence of the Strong View and the Normative Epistemic Claim. I do this by using a revised version of the Argument from Introspection that I deployed in the previous chapter. As in its original form, a key premise in this revised argument will be a claim about what we learn when we introspect our own suffering. I shall introduce the notion of impersonal introspection and will argue that when we impersonally introspect our own suffering, we know that we have a strong purely phenomenal reason to alleviate that suffering. Using this premise, I then show that when we empathize with the suffering of another person, we know the strength of our purely phenomenal reason to alleviate the suffering of that person.

Finally, in section 5, I return to the discussion of empathy-induced altruism that I began in chapter 2. In that chapter, I argued that the best explanation for empathy-induced altruism was the Phenomenal Epistemic Explanation, which says that empathy for a suffering person increases our altruistic motivation to help that person by giving us phenomenal knowledge of what that person's suffering is like. I will explore how either the Normative Epistemic Claim or the Strong View might be used to support the Phenomenal Epistemic Explanation.

As I have done in earlier chapters of this thesis, I will continue to focus my discussion on the question of what we come to know when we empathize with the

suffering of other people. For the purposes of this chapter in particular, I will narrow this focus further by considering what we come to know when we empathize with the undeserved suffering of other people. I do this to avoid overcomplicating the discussion with the issue of how considerations of desert impact our reasons to help other people who are suffering.¹²³

It is important to clarify that although I focus on the role that empathy has in informing us about our reasons to alleviate the present suffering of other people, I do not think that this is the only way that empathy can contribute to moral deliberation. At the end of the next, and final, chapter, I show that the epistemic power of empathy can also be utilized to give us knowledge of the reasons that we have to avoid harming other people, and to give us knowledge of the reasons that we have to promote joy in other people.

2. Reasons to Alleviate Suffering

There are, unfortunately, a wide range of ways that people can suffer. Physical suffering might, for example, arise through illness (ranging from common colds to cancer), injury (such broken bones), or hunger and malnutrition. Varieties of emotional suffering include depression, grief and loss, self-criticism, unexpressed frustration, anxiety and loneliness. Suffering can vary in intensity. In

¹²³ In defending the Strong View, I am committing myself to the position that we have strong reasons to alleviate any incidence of suffering, regardless of how it has come about and regardless of the moral status of the person who is experiencing the suffering. Accordingly, I think that if a criminal is suffering because of their 'deserved' legal punishment, then we have reasons to alleviate their suffering. This does not mean that, all things considered, we should not punish criminals, because there are important countervailing reasons for such punishment (e.g. the importance of protecting other members of society). I allow that we might have additional reasons to alleviate suffering when it is deserved, but I do not think that we can learn about such reasons through empathy (empathy with a suffering person cannot tell you whether their suffering is deserved).

its milder forms it may barely register in our awareness (e.g. a sore toe, or a mild unease about an upcoming event). In its stronger forms (e.g. the suffering involved during and after torture), it may be enough to make a life no longer worth living.

Sometimes, we have opportunities to alleviate our own suffering (e.g. taking a painkiller for a headache). On other occasions, we may be able to alleviate the suffering of a friend or loved one (e.g. by giving them support and reassurance when they are feeling distressed about a work assignment). The mechanics of modern charity mean that we are constantly being presented with opportunities to alleviate the suffering of strangers (e.g. by donating money to a charity which fights the spread of malaria). Where a person is suffering, we can engage in a practice of moral deliberation to assess the strength of reasons that we have to alleviate that person's suffering.

Most people would, I suspect, agree that if another person is suffering, we have a reason to help them.¹²⁴ This does not mean that alleviating a particular person's suffering is always what we have most reason to do, because we may have stronger reasons to perform an alternative action (such as helping a different person, or prioritizing our own projects and commitments). I also suspect that most people would agree that at least some of the reasons that we have to alleviate the suffering of other people are non-instrumental other-regarding reasons. To say that we have an other-regarding reason to alleviate another

¹²⁴ Of course, there may be exceptions. A nihilist would deny that we have any reason to do anything. An egoist would say that he only has reasons to alleviate his own suffering (although he might still take himself to have self-regarding reasons to alleviate the suffering of other people, where doing so will help him). As I discuss below, I think that both of these views are false. Moreover, I think that if a person who held either of these two views empathized with the suffering of another person, they would realize that their view was false because they would come to know that they had an other-regarding reason to alleviate the other person's suffering.

person's suffering is to say that we have a reason to help the other person for their own sake, and not only because we hope to derive some personal benefit from helping them.¹²⁵ To say that we have a non-instrumental reason to alleviate the suffering of another person is to say that we have a reason to alleviate their suffering that does not depend on the consequences that their suffering has. For example, depression may have bad consequences for a person by impeding their ability to accomplish important projects. This consequence of depression would give us an instrumental reason to help a depressed friend. Even if depression did not have bad consequences, however, we would still have a non-instrumental reason to help a depressed friend, because depression feels bad.

If it is true that we have non-instrumental other-regarding reasons to alleviate the suffering of other people, what else can we say about these reasons? I am inclined to think that agreement might be found for the claim that these reasons are proportional. To say that these reasons are proportional is to say that the strength of our reason to alleviate an incidence of suffering is proportionate to the severity of the suffering involved. For example, if one friend is heartbroken following a divorce and is struggling to leave the house, whilst another friend is feeling mild loneliness because his partner is away for the weekend, we have (other things equal) greater reason to help the former friend.

A more divisive question is this: how does the strength of our reason to alleviate suffering vary according to how we relate to the person who is suffering? A common view is that we have strong reasons to alleviate the suffering of friends

¹²⁵ Examples of ways that we might personally benefit from helping another person include: improving our social reputation, and increasing the chance that we receive reciprocal assistance in the future.

and family, but much weaker reasons to alleviate the suffering of strangers. We could call this a *Partial View*. More generally, this view would say that the closer we are to a person, the stronger our reasons to help them will be. Different versions of this view could be formulated to specify exactly what it means to be 'close' to another person, and to classify what different tiers of connection might be (for example, we could ask whether I am, by definition, 'closer' to a fellow citizen than I am to someone from a foreign country). An alternative view would be to say that we have strong reasons to alleviate all suffering, no matter who is experiencing it. We could call this an *Impartial View*. This view would say that the strength of reason that we have to eliminate an incidence of suffering does not depend on how we relate to the person who is suffering.

With this preliminary discussion in hand, I can now introduce the notion of a *purely phenomenal reason*. A purely phenomenal reason is a type of non-instrumental other-regarding reason. More specifically, it is the reason that we have to alleviate an incidence of suffering that is given to us solely by the phenomenal quality of the suffering (in other words, by what it is like). Where we are responding to a purely phenomenal reason, we alleviate an incidence of suffering simply because it feels awful. Purely phenomenal reasons are stronger for more severe instances of suffering because those instances of suffering feel worse. Using the terminology introduced above, we can say that purely phenomenal reasons are proportional.

By definition, the strength of a purely phenomenal reason to eradicate an incidence of suffering depends only on the phenomenology of the suffering, and not on any other fact about the suffering. For example, it does not depend on our

relation to the person that is experiencing the suffering.¹²⁶ If we have a purely phenomenal reason to alleviate the suffering of a sick relative, then, if a stranger is suffering in the same way, we will also have a purely phenomenal reason of the same strength to alleviate the suffering of the stranger. Similarly, if I have a purely phenomenal reason to alleviate my own suffering when I am feeling stressed, then I will also have a purely phenomenal reason of the same strength to alleviate the suffering of any other person who is feeling similarly stressed. To put the same point in a different way, purely phenomenal reasons are impartial.

In section 3 below, I argue that we have purely phenomenal reasons to alleviate suffering, and that these reasons are strong. It is important to note, however, that I will not be arguing that purely phenomenal reasons are the only kinds of reason that we have to alleviate suffering. Rather, I will allow that we may have many other kinds of reason to alleviate suffering. To illustrate this point, consider the following example. Suppose that two people are suffering from the same kind of headache, such that their experience is identical. Suppose also that the first person is a close friend of mine, and that the second person is a stranger. In this example, I will have purely phenomenal reasons to alleviate both incidences of suffering, and these reasons will be of the same strength in both cases. However, in each case I will also have other reasons to alleviate the suffering in question, and the nature and strength of these reasons may differ in each case. First, I may have self-interested reasons to help each person (e.g. if it will make me feel good to do so), and these reasons might be stronger in one

¹²⁶ The strength of a purely phenomenal reason to alleviate an incidence of suffering also does not depend on facts about how the suffering came about. Similarly, it does not depend on what consequences the suffering will have. This latter point explains why pure phenomenal reasons are a kind of non-instrumental reason.

case than in other (e.g. I might benefit more by helping my friend because he will be better company for me if he does not have a headache). Secondly, I could also have other-regarding instrumental reasons to help each person (e.g. the stranger could be a musician who will not be able to perform unless she gets rid of her headache). Thirdly, even when focusing on how bad the suffering feels for the other person, I might have a special reason to help my friend that arises from the fact that he is my friend (and not a stranger). This reason would be a non-instrumental other-regarding reason, but it would not be a purely phenomenal reason.

In the next section, I argue for the claim that we have strong purely phenomenal reasons to alleviate suffering. It is important to note that this claim does not commit me to a full-blown impartiality about our reasons to alleviate suffering. Whilst purely phenomenal reasons are themselves impartial, I allow that we can have many other types of reasons to alleviate the suffering of other people. As illustrated above, these 'other types of reason' may vary in strength depending on our relation to the person who is suffering. In other words, they can be partial.

3. Empathy as a Normative Guide

3.1 The Strong View

Having defined what I mean by a purely phenomenal reason, I can now explain the sense in which I take empathy to be an important normative guide. In this section, I will make two claims. The first is that we have strong purely phenomenal reasons to alleviate suffering. The second is that empathy is a way

of coming to know the strength of these reasons. I begin with the first claim and I call it the *Strong View*.

The Strong View: If a person B is experiencing an incidence of suffering S, then, for any other person A, A has a strong purely phenomenal reason to alleviate S.

It is hard to specify precisely what I mean in saying that purely phenomenal reasons are strong. I offer four clarificatory remarks here.

First, I do not mean to contradict the point made above that purely phenomenal reasons are proportional (to the severity of the suffering involved). Accordingly, I allow that where an incidence of suffering is very mild, our purely phenomenal reasons to alleviate that suffering will be weak. Secondly, in saying that purely phenomenal reasons are (in general) strong, I mean that they are stronger than we might think. In other words, I mean that they are stronger than conventional morality would lead us to believe. As I shall explain further below, my view is that although we generally take ourselves to have purely phenomenal reasons to alleviate the suffering of other people, we are typically surprised when we find out (through empathy) how strong those reasons are. In short, my view is that we typically understate the strength of our purely phenomenal reasons to alleviate suffering. Thirdly, in saying that purely phenomenal reasons are strong, I mean that they should be an important consideration in decision-making. Even if such reasons can be outweighed by other considerations, they should be taken seriously. Finally, in saying that purely phenomenal reasons are strong, I mean that the purely phenomenal reasons that we have to alleviate the suffering of

other people are just as strong as the purely phenomenal reasons that we have to alleviate our own suffering.

3.2 The Normative Epistemic Claim

The second claim that I defend in this chapter says that empathy is a way of learning about the strength of our purely phenomenal reasons to alleviate the suffering of other people. I call this claim the *Normative Epistemic Claim*.

The Normative Epistemic Claim: When a person A empathizes with a suffering person B who is feeling S (an incidence of B's suffering), then A's empathy for B will give A knowledge (if A did not have it already) of the strength of the purely phenomenal reason that A has to alleviate S.

To illustrate this claim, suppose that I have a friend who has just been in a car crash, and who needs to spend a prolonged period resting alone at home. In addition to her physical pain, my friend is also suffering from loneliness. I might empathize with this friend by consciously taking up her perspective, or through affective contagion, or perhaps through remembering my own similar experiences. The Normative Epistemic Claim says that if I do so, then I will know the strength of my purely phenomenal reason to alleviate my friend's suffering (I might, for example, try to help my friend by visiting her for a few hours). The purely phenomenal reason that I have to help my friend is a reason that is given to me solely by the phenomenal quality of her suffering. When I act for that reason, I will say that I am helping my friend because her suffering feels awful. I may also have additional reasons to help my friend (e.g. because her suffering will prevent her from passing an important exam, or because she is my friend).

The Normative Epistemic Claim does not say that empathy will teach me about these reasons.

Given what I have said about empathy in previous chapters, one can see why it is ideally suited to learning about purely phenomenal reasons. Empathy, I have said, is the means through which we acquire phenomenal knowledge about the suffering of other people. By definition, a purely phenomenal reason is the reason that we have to alleviate an incidence of suffering that is given to us solely by the phenomenal quality of that suffering. Consequently, knowing about the phenomenology of an incidence of suffering seems to be an obvious first step in coming to know about one's purely phenomenal reason to alleviate that suffering.

It is important to note that although the Normative Epistemic Claim says that empathy is one way of acquiring knowledge about the strength of our purely phenomenal reasons to alleviate suffering, it does not claim that empathy is the only way of acquiring this knowledge. Perhaps we could also acquire this knowledge by listening to a trusted moral expert. Alternatively, a guide might be created which, for every possible form of suffering, listed the strength of our purely phenomenal reason to alleviate that suffering.

It should be noted, however, that merely knowing that another person is suffering (even if this involves knowing the type and severity of their suffering) does not mean that we will know the strength of the purely phenomenal reason that we have to help that person.¹²⁷ First, this is because we might know that a person is suffering, but fail to consider what reason we have to help them. We

¹²⁷ For a related discussion, see section 2 in the previous chapter. In that section, I identified several types of case in which we can know which affective state another person is feeling but fail to know the intrinsic value of that affective state to that person.

might fail to consider this because, for example, we are concerned about how their suffering will impact on us, or because we are focused on trying to work out who is responsible for causing their suffering. Secondly, we may consider the strength of the reason that we have to help the other person but assess the strength of that reason incorrectly. One way in which we could do this is by understating the strength of that reason.¹²⁸ We might do this because we underestimate the impact of the other person's suffering on their well-being, or because we fail to recognize the extent to which we have reason to help the other person when their well-being is low.

Indeed, according to the Strong View, we do typically underestimate the strength of our purely phenomenal reasons to help other people. Even if we could come to the correct view about the strength of these reasons without the use of empathy, the Strong View says that we do not. The implication is that when we empathize with the suffering of another person, our empathy for that person will typically lead to an upward revision in our judgement about the strength of our purely phenomenal reason to help that person. In other words, we can say that empathy has a 'radical educational function' in so far as it makes us realize that the suffering of others matters far more than we think it does.

Empathy can have such a function even with altruistically minded people who already take themselves to have reasons to alleviate the suffering of other people. It is perhaps for this reason that charities try to induce potential donors to empathize with the people that they could help. Moreover, the idea that empathy

¹²⁸ It is also possible that we could overstate the strength of the purely phenomenal reason that we have to help another person. We might do this, for example, if we think that their suffering is worse than it actually is.

has such a function can be used to explain examples of people who dedicate their lives to helping disadvantaged groups of people after vividly and deeply empathizing with members of that group.¹²⁹

If the Strong View and the Normative Epistemic Claim are true, then empathy can also lead to even more fundamental conversions in a person's normative outlook. For example, suppose that an egoist (who believed he had no other-regarding reasons) empathized with the suffering of another person. If the aforementioned claims are correct, then the egoist will come to know that he has a strong other-regarding reason to alleviate the other person's suffering. Accordingly, her empathy will demonstrate to her that her egoism was false. In this vein, a parent might ask a self-centred child to empathize with the feelings of another child that they are teasing (the parent might say, 'How would you feel, if you were them?').

Moreover, on the approach that I am outlining here, empathy could also be used to educate a person out of holding discriminatory moral views. Consider the example of an extreme racist who believed that certain racial groups were morally irrelevant, such that he thought that he had no reason to alleviate their suffering. If such a racist empathized with the suffering of someone within a racial group that he deemed to be 'irrelevant', he would learn that his prior moral views were false. More specifically, he would learn that he had a strong purely phenomenal reason to alleviate the suffering of the person that he empathized with. One would hope that he would also be able to generalize this learning, by realizing that he had

¹²⁹ Psychologist Martin Hoffman has a fascinating discussion of examples of this sort, which he calls examples of 'witnessing' (Hoffman, 2011: pp. 238-245).

similar reasons in relation to any other member of that racial group who was suffering.

3.3 The limitations of empathy

It is important to clarify that the Normative Epistemic Claim does not say that a single act of empathy will give the empathizer knowledge of what they have most reason to do in a given situation. Instead, the Normative Epistemic Claim says only that an act of empathy will give the empathizer knowledge of one of their reasons for action (namely, their purely phenomenal reason to alleviate the suffering of the person that they are empathizing with). There is a range of other kinds of reason that the empathizer might not know about.

First, the empathizer might not know anything about the purely phenomenal reasons that they have to alleviate the suffering of people that they have not empathized with. Even if they could acquire that knowledge through empathizing with additional people, they may not do so.

Secondly, the empathizer might lack knowledge about their reasons for action that are not purely phenomenal reasons. They may have self-regarding reasons for helping, or not helping, the person that they are empathizing with. They may also have instrumental other-regarding reasons for helping that person (e.g. they might have reason to help alleviate a friend's back pain because the friend wants to run a marathon), or for not helping them (e.g. perhaps helping a stressed friend will encourage an unhelpful pattern of dependency). The empathizer might also have special reasons to help a person because they are friends with that person, or because they are closely related to that person.

Finally, the empathizer might have reasons for action that are not inherently tied to the alleviation of suffering. For example, they could have reasons to follow moral rules such as 'always keep promises' or 'never steal'. If an empathizer should not help the person that they are empathizing with because they have already promised to help someone else, this is not a fact that their empathy will reveal to them.

It should be clear, therefore, that it is consistent with the Normative Epistemic Claim for an empathizer to have a false view about what they have most reason to do. In the first chapter, I discussed the Spotlight Effect Objection to empathy which said that empathy can lead the empathizer to falsely believe that they should help the person that they are empathizing with (when they should in fact help someone who is in greater need, or who is easier to help). In the next (and final) chapter, I discuss how best to respond to this objection. For now, it should simply be noted that the Normative Epistemic Claim is compatible with this objection.

As a final point in this section, it is important to note one further limitation on the epistemic power of empathy: it does not tell us how useful any particular helping action will be. According to the Normative Epistemic Claim, empathy tells us only about the strength of reason that we have to help a person. It does not tell us which actions (if any) will actually be helpful. We might know that we have a reason to help a depressed friend but be unsure of how to help them. Similarly, we might know that we have a reason to help a malnourished child living in a far-off country but fail to know how we could effectively (and sustainably) do so.

Having outlined the Strong View and the Normative Epistemic Claim, I now proceed to defend both of these claims using a revised version of the Argument from Introspection.

4. The Revised Argument from Introspection

4.1 Recap

The purpose of this section is to defend the Strong View and the Normative Epistemic Claim. The former says that where there is suffering, we have a strong purely phenomenal reason to alleviate it, regardless of who it is that is suffering. The latter says that when we empathize with a person who is suffering, our empathy for that person will give us knowledge of the strength of our purely phenomenal reason to help them. The argument that I deploy here is similar to that which I offered in the previous chapter, in so far as it begins with a claim about what we learn when we introspect our own suffering. Accordingly, it will be helpful if I begin this section by recapping the argument from the previous chapter.

I began that argument by claiming that when we introspect an incidence of our own suffering, we know how intrinsically bad it is for us. We also know, I continued, that our suffering is intrinsically bad for us because of how it feels, such that any other incidence of suffering that feels the same way will be just as bad for whoever is experiencing it. The implication is that if we introspect an incidence of our own suffering in order to empathize with another person who is suffering in a similar way, then we will know: (i) how bad our suffering is for us, (ii) that the other person's suffering feels just the same as ours, and therefore (iii)

that it is just as bad for them as our suffering is for us, and thus (iv) how bad their suffering is for them. This was the original version of the Argument from Introspection that I used to defend the Prudential Epistemic Claim.

This version of the Argument from Introspection cannot be used to establish the Normative Epistemic Claim or the Strong View. To see this, note that all of the statements made in this version of the argument could be true whilst it was also true that we had no other-regarding reasons whatsoever to alleviate the suffering of other people. This could be because we have no other-regarding reasons to promote the well-being of other people. Moreover, even if we do have other-regarding reasons to alleviate the suffering of other people, the Argument from Introspection does not show that an empathizer will know this. Instead, the empathizer could falsely believe that he had no reason to promote the well-being of the person that he was empathizing with, such that he could know that their suffering was bad for them but (falsely) believe that he had no reason to help them.

Accordingly, to defend the Normative Epistemic Claim and the Strong View, I now develop a revised version of the Argument from Introspection. The crucial modification concerns what we learn when we introspect our suffering. Instead of saying that introspection gives us knowledge of how intrinsically bad our suffering is for us, the revised argument says that introspection gives us knowledge of the strength of the purely phenomenal reasons that we have to alleviate our suffering. To be more specific, the revised argument says that it is impersonal introspection that gives us this knowledge.

I develop this argument in three stages. First, I argue that introspecting our own suffering gives us knowledge of the reasons that we have to alleviate it and that these reasons are strong. Secondly, I introduce the notion of impersonal introspection and argue both (i) that impersonally introspecting our own suffering gives us knowledge of the strength of the purely phenomenal reason we have to alleviate it, and (ii) that the use of such impersonal introspection shows that our purely phenomenal reasons to alleviate suffering are strong. Finally, I argue that empathy involves impersonal introspection and thus that it gives us knowledge of the strength of the purely phenomenal reason that we have to alleviate the suffering of the person that we are empathizing with.

4.2 Introspection and Reasons

Suppose that I am suffering from severe back pain. In the previous chapter, I claimed that if I introspect my back pain, I will come to know how intrinsically bad it is for me. In other words, I will know the extent to which it makes my life worse, independently of any consequences that it might have. The further claim that I now wish to make is that if I introspect my back pain, I will also know that I have non-instrumental reasons to alleviate it and I will know how strong these reasons are.¹³⁰ Indeed, through introspecting my back pain, I will know that these reasons to alleviate my back-pain are strong (in the sense that they are not easily outweighed by other considerations). More generally, I suggest that, whenever we are experiencing suffering, we have strong non-instrumental

¹³⁰ I may also have instrumental reasons to alleviate my back pain. For example, if persistent back pain will stop me from sleeping, then this gives me an additional (instrumental) reason to alleviate my back pain. I am not saying that introspection will give us knowledge of our instrumental reasons to alleviate our suffering.

reasons to alleviate that suffering, and we can know the strength of these reasons through introspecting our suffering.

Two clarifications are important here. First, in saying that we have strong non-instrumental reasons to alleviate our own suffering, I do not just mean that we want to alleviate our own suffering. Similarly, in saying that we learn about our non-instrumental reasons to alleviate our own suffering through introspection, I do not just mean that we learn how much we want to alleviate our own suffering. Rather, I am claiming that it is rationally appropriate to alleviate our suffering and that we learn this fact whenever we introspect an incidence of our suffering. A person who had no motivation to alleviate her own suffering would be failing to be motivated to do what she had reason to do.¹³¹

Second, in saying that a person who introspects their suffering will know the strength of their non-instrumental reason to alleviate that suffering, I mean (amongst other things) that they will not be uncertain about this reason. In other words, this person will not say, 'Yes I know that my suffering is bad for me, but do I have a reason to alleviate it?' When we are not introspecting our suffering, such doubt may be entertained. One might, in the abstract, question whether we have any kind of reason to promote our own well-being. When we are introspecting our suffering, however, we know that we have a non-instrumental reason to alleviate it such that there is no room for doubt. To test this claim one must, unfortunately, suffer (and introspect one's suffering).

¹³¹ Although we always have reason to alleviate our suffering, I am not saying that we always have reason to alleviate our pain. As discussed in the previous chapter, if pain is not disliked then it may not be bad for us and we may have no reason to alleviate it.

If introspection of our suffering gives us knowledge of the non-instrumental reasons that we have to alleviate it, then two interesting questions arise: how does introspection give us this knowledge? And: how does this knowledge relate to the knowledge that our own suffering is intrinsically bad for us? I do not mean to answer these questions conclusively here, but I do wish to briefly suggest three possible answers. First, we might acquire the two pieces of knowledge separately. This answer says that although both pieces of knowledge are acquired through introspection, neither piece of knowledge is derived from the other. Secondly, we might first know the badness of our suffering and then utilize this knowledge to acquire the further knowledge that we have non-instrumental reasons to alleviate our suffering. Thirdly, we might first know that we have non-instrumental reasons to alleviate our suffering, and then utilize this knowledge to acquire the further knowledge that the suffering is bad for us. Whichever of these routes one chooses, one's choice does not commit one to a position in the 'buck-passing' debate. In other words, whatever one's view is about the epistemology of how we learn about reasons and value, it does not commit one to a view about the ontological relationship between reasons and value. For example, even if we learn about our reason to alleviate our own suffering through first learning about the badness of our own suffering, it could still be that the latter was explicable in terms of the former.¹³²

¹³² The following analogy may be helpful here: events in a computer's hardware determine what it displays on its monitor, but we may learn about the former by observing the latter. In the same way, we might learn about our non-instrumental reasons to alleviate suffering by first learning about the badness of suffering, even if the badness of suffering is itself explained by the reasons we have to alleviate suffering.

4.3 Impersonal introspection and reasons

Earlier, I introduced the notion of a purely phenomenal reason to alleviate suffering. By definition, the strength of the purely phenomenal reason that we have to alleviate an incidence of suffering depends only on the phenomenology of that incidence of suffering and not on any other fact about the suffering. Consequently, the strength of such a reason does not depend on how one relates to the person who is suffering. If I have a strong purely phenomenal reason to alleviate my own suffering, then the fact that this suffering is mine plays no role in determining the strength of this reason. I would have a reason of the same strength to alleviate any other incidence of suffering that was phenomenally identical to mine.

The claim that I now wish to defend is that there is a way of introspecting our own suffering which gives us knowledge that we have purely phenomenal reasons to alleviate our own suffering and which gives us knowledge that these reasons are strong. I shall refer to this 'way of introspection' as *impersonal introspection*.

In impersonal introspection, we introspect our suffering and relate to it not as our suffering, but simply as suffering. We focus entirely on the phenomenology of the suffering and let go of any sense that it is ours. When we do this, I claim, we know not only that we have strong non-instrumental reasons to alleviate our suffering, but also that we have strong purely phenomenal reasons to alleviate our suffering.¹³³ When we impersonally introspect our suffering, we find the

¹³³ In other words, I am saying that purely phenomenal reasons are an important type of the non-instrumental reasons we have to alleviate our suffering. I am not saying, however, that the purely phenomenal reasons that we have to alleviate our own suffering are the only kind of non-instrumental reason that we have to alleviate our own suffering. We may also have self-regarding

identity of the sufferer to be irrelevant. What matters is solely the phenomenology of the suffering (that is, how it feels). If we entertain the possibility that another person is experiencing an incidence of suffering that feels just like ours, it will be clear to us that we have a strong non-instrumental reason to alleviate their suffering. This will be true no matter who the other person is because what matters is only what their suffering feels like.

The perspective of impersonal introspection that I am describing here is similar to a perspective described by Thomas Nagel in *A View From Nowhere* (Nagel, 1986). In a section in which he argues that we have ‘agent-neutral reasons’ to eradicate the pain of others, Nagel describes taking up an objective perspective on our subjective mental states.¹³⁴ In adopting this perspective, we view our suffering both from the inside, because that is the only way of capturing its phenomenology, and also from the outside, in so far as we do not view it as ours. Nagel says that when we take up this perspective with respect to an experience of pain, we know that we have reason to eradicate any experience that feels like that:

But the pain, though it comes attached to a person and his individual perspective, is just as clearly hateful to the objective self as to the subjective individual. I know what it’s like even when I contemplate myself from outside, as one person amongst countless others... The pain can be detached in thought from the fact that it is mine without losing any of its dreadfulness. (Nagel, 1986: p. 160)

non-instrumental reasons to alleviate own suffering that we do not have in relation to the suffering of other people.

¹³⁴ Although Nagel talks about pain, and not about suffering, his cases of pain are cases of suffering (disliked pain) in so far as he thinks that pain is something that we are strongly disposed to dislike. It should be noted that whereas I think dislike makes suffering bad by making it feel bad, Nagel does not seem to hold this view. Instead, he holds what I called the Dislike View about the badness of pain (see chapter 4). Consequently, on Nagel’s view, when we take up an objective perspective on another person’s pain by empathizing with them, we take ourselves to have a reason to alleviate that pain because we endorse the subject’s dislike of their pain. My view, by contrast, is that we judge ourselves to have a reason to alleviate their pain because we observe how the subject’s dislike of their pain makes their suffering feel awful.

In saying that pain is 'dreadful', Nagel means that it is 'objectively bad' and that we thus have 'agent-neutral' reasons to alleviate it. Such a reason applies to every agent, regardless of their personal desires and projects, and regardless of their relation to the person who is in pain. To put Nagel's point into my terminology, Nagel argues that if we impersonally introspect an incidence of our pain, we know that we have a purely phenomenal reason to eradicate it, such that we would have the same reason to eradicate any other incidence of pain that felt the same way. In short, we will think, '*This experience* ought not to go on, whoever is having it' (ibid.: p. 161). I now move to show that this is precisely the thought that arises for an empathizer, because empathy involves impersonal introspection.

4.4 The applications to empathy

Above, I argued that if we impersonally introspect an incidence of our suffering, we will know that we have a strong purely phenomenal reason to alleviate it. If I can show that empathy involves such impersonal introspection, then I will have done enough to establish the Normative Epistemic Claim. To see this, consider a brief example. Suppose that a friend of mine is suffering with anxiety about an upcoming interview and that I empathize with this friend by vividly recalling a similar experience, or by imaginatively taking up his perspective. As part of this empathy, I will thus come to feel an incidence of anxiety that is similar to that felt by my friend. By the argument above, if I impersonally introspect my anxiety (which is an incidence of suffering), I will know that I have a strong purely phenomenal reason to alleviate it. When I consider that

my friend's anxiety feels just the same as mine, I will thus know that I have the same strength of purely phenomenal reason to alleviate his anxiety. The fact that his suffering is his and that my suffering is mine is irrelevant. When it comes to purely phenomenal reasons to alleviate suffering, what matters is only the phenomenology of suffering, and his suffering feels just like mine. More generally, we can say that when I empathize with a suffering person, I will come to know the strength of my purely phenomenal reason to alleviate that person's suffering. This is the Normative Epistemic Claim.

However, at this point, one may ask: why should we think that empathy involves impersonal introspection? When I defined empathy in chapter 3, I said that it necessarily involves introspection, but I did not say that it necessarily involves impersonal introspection. Nevertheless, I now argue, empathy as I defined it does necessarily involve impersonal introspection.

My account of empathy requires that the empathizer, in introspecting their empathic affect (which is similar to that felt by the other person), actively considers the belief that the other person is having an experience just 'like this'. In other words, the empathizer utilizes their own affect as a way of representing what the other person is experiencing. To do this, they must focus on what is common between the two affective states and abstract away from what is different between them. What is common between the two affective states is their phenomenology and what is different between them is who is experiencing them. Consequently, the empathizer views their affective state not as theirs, but rather as an instance of an affective state that could, in general, be experienced by anybody. In any particular act of empathy, the empathizer frames their affective state as a representation of the affective state that is being experienced by the

person that they are empathizing with. In short, the empathizer necessarily takes an impersonal stance towards their own affective state. I am calling this stance impersonal introspection. By the argument above, when the empathizer engages in such introspection, they will know that they have a strong purely phenomenal reason not just to alleviate their own suffering, but to alleviate any suffering that feels just like theirs. Consequently, they will know that they have a strong purely phenomenal reason to alleviate the suffering of the person that they are empathizing with.

5. The Phenomenal Epistemic Explanation Revisited

So far in this chapter, I have been arguing for the Strong View and for the Phenomenal Epistemic Claim. As a final task in this chapter, I now wish to discuss briefly how those two claims might be used to strengthen the Phenomenal Epistemic Explanation of empathy-induced altruism that I outlined in chapter 2 of this thesis.

As I pointed out in that chapter, there is considerable evidence in favour of the Empathy-Altruism Hypothesis.

The Empathy – Altruism Hypothesis: If a person A empathizes with a suffering person (B), there will normally be an increase in A's altruistic motivation to help B.¹³⁵

¹³⁵ As I noted in chapter 2, one of my aims in explaining how empathy gives rise to altruistic motivation is to help specify the conditions under which it will give rise to such motivation. This would allow us to explain what is meant by the term 'normally' in the Empathy-Altruism Hypothesis.

I also argued that there is an as-yet-unanswered question about how to explain this hypothesis. In other words, I said that we lack a satisfactory explanation of how empathy for a suffering person gives rise to an increase in altruistic motivation to help that person. I concluded by outlining a new explanation of empathy-induced altruism: the Phenomenal Epistemic Explanation. This explanation says that empathy for a suffering person gives rise to an increase in altruistic motivation to help that person by giving the empathizer new phenomenal knowledge about what that person's suffering is like.

My primary goal in introducing this explanation was to help develop my case that empathy plays an important role in moral deliberation. I sought to give credence to the idea that empathy (and empathy alone) is a source of phenomenal knowledge about the affective states of other people. I called this idea the Strong Phenomenal Epistemic Claim and I defended it in chapter 3. In this chapter, and in the previous chapter, I have then progressed to show that empathy, by giving us phenomenal knowledge about the painful affective states of other people, also gives us knowledge about the intrinsic prudential badness of those affective states and about our purely phenomenal reasons to alleviate those states. Accordingly, I have shown that empathy is an important input into moral deliberation. I shall return to this point in the next chapter.

As I said above, what I now wish to do is to show how the claims of this chapter might be used to strengthen the Phenomenal Epistemic Explanation. In particular, I wish to show how they can be used to support the idea that acquiring phenomenal knowledge about the suffering of another person increases our motivation to alleviate that suffering.

To begin, a warning is required: the issue of how reasons for action relate to motivation is a complex and divisive one, and I cannot do it full justice here. Nevertheless, I wish to outline two approaches that might be taken to explaining how phenomenal knowledge motivates an empathizer.

The first approach takes as its starting point an Internalist theory of normative reasons, of the sort developed by Williams (1979).

Internalism About Reasons: For it to be true that (i) a fact F about an action X gives person A a reason to do X, it must be true that (ii) if A knew F, and if A deliberated rationally, then A would be motivated to do X.¹³⁶

We can combine this theory with the Strong View, which says that we have strong purely phenomenal reasons to alleviate suffering. The Strong View implies that if one of my actions will alleviate another person's suffering, then I have a strong reason to do that action. Moreover, it says that this reason is given to me solely by the fact of what the other person's suffering is like (it is because their suffering feels awful that I have a purely phenomenal reason to alleviate it). According to Internalism About Reasons, it follows that if I knew what the other person's suffering was like, and if I deliberated rationally, then I would be motivated to do any action that I thought would alleviate their suffering. Thus, on the assumption that I deliberate rationally, acquiring phenomenal knowledge about the other person's suffering will increase my motivation to help them.

¹³⁶ My interpretation of Williams' position has been influenced by reading Parfit (1997). Roughly speaking, to deliberate rationally is to seek to make one's beliefs, desires and motivations consistent with one another. It does not involve having any particular desires or commitments. The rationality that I am talking about here is thus procedural rationality (which is to be contrasted with substantive rationality).

The reasoning above demonstrates that if both Internalism About Reasons and the Strong View are true, then phenomenal knowledge would be motivating.¹³⁷ Because this motivation arises from our disposition to eradicate suffering in whomever it arises, it seems a plausible further step to say that this motivation is altruistic.¹³⁸ Accordingly, we could say that it is by giving us phenomenal knowledge that empathy gives rise to altruistic motivation.

It is interesting to note that on the above approach the normative knowledge that the empathizer acquires (of their purely phenomenal reasons for action) does not play any role in increasing their altruistic motivation. Instead, what gives rise to the increase in their altruistic motivation is the phenomenal knowledge about what the person they are empathizing with is feeling.

On a second approach to developing the Phenomenal Epistemic Explanation, we can suppose that it is the normative knowledge acquired by the empathizer that increases their motivation to help the person that they are empathizing with. This approach supposes that, generally speaking, moral agents are motivationally responsive to their normative beliefs, such that if they come to believe that they have a reason to perform an action, then there will be an increase in their motivation to perform that action. To put the same point in a different way, this approach says that agents are reason-responsive and that reason-responsiveness is part of what it means to be a rational agent.

¹³⁷ That is, on the assumption that the empathizer deliberates rationally. From now on, and for the sake of brevity, I shall adopt this assumption.

¹³⁸ This use of the term 'altruistic' may be inconsistent with the way that Batson defines altruism. As discussed in chapter 2, Batson says that altruistic motivation is motivation with the ultimate goal of promoting another person's well-being. Batson's definition may thus fail to include the motivation that I am talking about here, which has as its ultimate goal the eradication of suffering in whomever it arises.

On this approach, it is the Normative Epistemic Claim (rather than the Strong View) which plays a key role in supporting the Phenomenal Epistemic Explanation. According to the Normative Epistemic Claim, by empathizing with a suffering person the empathizer will come to know that they have a strong purely phenomenal reason to alleviate that person's suffering. If this person is reason-responsive, this knowledge will thus increase their motivation to help the person that they are empathizing with. In this way, the Normative Epistemic Claim could be used to explain empathy-induced altruism.¹³⁹

We might associate this approach with *Externalism About Reasons*, which is the theory that Internalism About Reasons is false. Externalism About Reasons says that a fact about an action can give me a reason to do that action even if knowing that fact, and deliberating rationally, would not motivate me to perform that action. However, it should be noted that Externalism About Reasons does not entail that humans are motivationally reason-responsive, because an Externalist is not committed to saying anything about our motivational responses to reasons (or about our motivational responses to our beliefs about reasons).

The two approaches outlined above can each help to develop our understanding of the circumstances under which exceptions to the Empathy-Altruism Hypothesis might arise. According to the Strong Phenomenal Epistemic Claim, empathy with a suffering person will necessarily give me phenomenal knowledge of what that person's suffering is like. Although, on either of the above

¹³⁹ As with the first approach, there is a question here about whether Batson would count this empathy-induced motivation to be altruistic motivation. Given that the motivation here arises from our general disposition to be motivated to do what we believe we have reason to do, one might suggest that its ultimate goal consists in 'living in accordance with reason' rather than in promoting the well-being of another person. Nevertheless, I suggest that this motivation still counts as being 'altruistic' on an everyday usage of that term (and it certainly is not what Batson would call egoistic motivation).

approaches, such knowledge will normally increase my altruistic motivation to help the other person, there are possible exception cases. On the first approach, I may fail to be motivationally responsive to phenomenal knowledge. In this scenario, I would thus also serve as an exception case to the Strong View because it could not be true that I had purely phenomenal reasons to alleviate suffering.¹⁴⁰ On the second approach, I may fail to be reason-responsive. Even though I had a purely phenomenal reason to help the other person, and even though I knew this, I may fail to be motivated by this knowledge.

Although these exception cases to the Empathy-Altruism Hypothesis could arise in theory, I am sceptical that they arise in practice. Psychopaths are the standard example of a type of person that is insensitive to the pain of others. Yet although psychopaths may know a great deal about what other people are feeling (they can acquire a lot of attributive knowledge), I doubt that they acquire phenomenal knowledge about what other people are feeling. I doubt this because I doubt that they empathize with other people.¹⁴¹ As I have defined it, empathy requires a considerable amount of emotional sensitivity and self-awareness. To empathize with the suffering of another person requires one to remember, access, or imagine a form of suffering that is similar to that felt by the other person. As far as I know, this is not something that psychopaths are likely to do. Accordingly, they do not serve as examples of people who acquire phenomenal

¹⁴⁰ I could not have such reasons because, on the first approach, Internalism About Reasons is true.

¹⁴¹ It is often suggested that the moral deficits of psychopaths can be explained in terms of their failure to empathize with other people. (Blair, 1995; Nichols 2004). Moreover, even if psychopaths might be better at sympathy or emotional contagion than previously thought (Lishner et al, 2012), that does not mean that they can empathize with other people.

knowledge about another person's suffering and yet fail to be moved to help that person.

6. Conclusion

In this chapter, I began by introducing the notion of a purely phenomenal reason to alleviate suffering. I then argued for the Strong View, which says that we have such reasons and that they are strong, and for the Normative Epistemic Claim, which says that empathy for a suffering person gives us knowledge of our purely phenomenal reasons to help that person. To argue for these claims, I introduced the notion of impersonal introspection. I said that if we introspect an incidence of our suffering and relate to it simply as suffering, rather than as our suffering, we will know that we have a strong purely phenomenal reason to alleviate it. At the end of this chapter, I explored how either the Strong View or the Normative Epistemic Claim might be used to strengthen the Phenomenal Epistemic Explanation of empathy-induced altruism that I introduced in chapter 2.

My overall goal in this thesis is to vindicate the role of empathy in moral deliberation. This chapter has taken an important step towards achieving that goal by showing that empathy provides us with knowledge about an important kind of reason for action. In the next, and final, chapter, I will summarize the importance of empathy to moral deliberation. I shall also show how the Normative Epistemic Claim can be used to respond to the Spotlight Effect Objection to empathy that I discussed in the opening chapter of this thesis. Finally, I will explore how empathy can be used not just to respond to the present suffering of other people, but also

to learn about our reasons to not harm people, and to learn about our reasons to bring joy to other people.

Chapter 6: Empathy and Moral Deliberation

1. Overview

In this chapter, I draw together the insights of the five preceding chapters to summarize my account of how empathy contributes to moral deliberation. In developing this account, my aim has been to strengthen the Epistemic Defence of empathy. The standard version of that defence claims that empathy is important to moral deliberation because it provides us with attributive knowledge about the affective states of other people. In the next section of this chapter (section 2), I recap how I have built upon this claim by showing that empathy gives rise to additional types of knowledge about the affective states of other people.

Following that, I revisit (in section 3) the point made by critics of empathy that empathy can lead our decision-making astray by making us have too much concern for the person that we are empathizing with. In the first chapter of this thesis, I called this point the Spotlight Effect Objection. Using the two claims that I defended in the previous chapter (the Normative Epistemic Claim and The Strong View) I show how the supporter of empathy can best respond to this objection.

So far in this thesis, I have focused on exploring what we learn when we empathize with the present suffering of other people. However, it is important to note that other kinds of empathy can also play a powerful role in moral deliberation. Consequently, in section 4, I consider the importance of empathizing with other people with respect to the harm that we could cause them. Finally, in section 5, I argue that empathizing with the joy of other people can also inform our moral deliberations, in so far as it gives us knowledge of the reasons that we have to help cultivate such joy.

2. The Epistemic Defence of Empathy

As I described it in the opening chapter of this thesis, the Epistemic Defence of empathy says that empathy is an important way of informing our moral deliberations. This defence does not say that empathy is necessary for making moral judgements, or even that it is necessary for making accurate moral judgements.¹⁴² Rather, this defence says that empathy can help us to make informed moral judgments by providing us with knowledge about the affective states of other people.

As noted above, I have so far been focusing my discussion on the importance of empathizing with the present suffering of other people. In other words, I have been discussing the importance of empathizing with people who are feeling painful affective states (emotions, moods and physical sensations) that they dislike.¹⁴³ Using the definition of empathy that I gave in chapter 3 (empathy as experiential understanding), we can say that such empathy involves (i) feeling an affective state that is similar to that felt by the other person, (ii) introspecting this 'similar' affective state, and (iii) forming a conscious and justified belief that the other person is feeling something 'like this'. One might come to feel the 'similar' affective state through affective contagion, through imaginatively taking up the other person's perspective, or through utilizing one's own life experiences.¹⁴⁴

¹⁴² Accordingly, it is consistent with Jesse Prinz's contention that empathy is not necessary for morality (Prinz 2011a and 2011b). For more detail on Prinz's view, see section 3.2 in chapter 1.

¹⁴³ In chapter 4, I defined suffering as the combination of a painful affective state and the consciously felt dislike of that state.

¹⁴⁴ I discussed the mechanisms of perspective-taking and affective contagion in chapter 1. In chapter 3, I introduced the idea that we can also utilize our own life experiences to empathize with other people.

According to the Epistemic Defence, such empathy is a way of acquiring knowledge that will help us to deliberate about when to alleviate the suffering of other people. However, it must be acknowledged that empathy comes with an emotional cost. Empathizing with the painful affective state of another person involves feeling a similar painful state that is similarly painful. In the first chapter, I referred to this point as the Emotional Cost Objection to empathy. That objection adds that the pain of empathizing with others can be draining and debilitating, to the point where it may even impede the empathizer's ability to help the person that they are empathizing with.

The standard version of the Epistemic Defence says that the epistemic contribution of empathy to moral deliberation is attributive knowledge. In other words, it says that empathy with a suffering person contributes to moral deliberation about whether to help that person by giving us knowledge of which affective states that person is feeling. This version of the Epistemic Defence is vulnerable to the criticism that testimony and inference provide alternative ways of acquiring attributive knowledge that are less emotionally costly. For instance, the critic of empathy can say, it is easier to ask someone in pain how they are feeling than it is to feel how they are feeling.

The standard version of the Epistemic Defence is also vulnerable to what I have called the Spotlight Effect Objection to empathy, which says that empathy can lead our decision-making astray by making us falsely believe that we should help the person that we are empathizing with. I shall return to consider that objection in the next section. For now, I wish to summarize how I have strengthened the Epistemic Defence by showing that empathy contributes more than attributive knowledge to moral deliberation.

Firstly, I have argued that empathy also contributes phenomenal knowledge to moral deliberation. Indeed, I argued for the claim that empathy is the only source of phenomenal knowledge about the affective states of other people. I called this claim the Strong Phenomenal Epistemic Claim. After introducing this claim in chapter 2 (as part of my proposed explanation for the Empathy-Altruism Hypothesis), I defended it in chapter 3. According to the Strong Phenomenal Epistemic Claim, empathy with a suffering person does not just give us knowledge of which affective states that person is feeling, but it also gives us knowledge of what it is like for that person to feel such affective states. Indeed, the Phenomenal Epistemic claim says that empathy is the only way of acquiring this knowledge.

Intuitively, phenomenal knowledge seems to be important to moral deliberation. Thus, if I knew that a friend was depressed but did not know what it was like for him to feel depressed, my intuition says that I would be missing something important in my deliberations about whether to help my friend. I would fail to know what my friend's depression was like if I had never experienced depression and (I have argued) I would also fail to know what it was like if I had experienced depression but had nevertheless failed to empathize with my friend. In chapters 4 and 5, I went on to defend this intuition that phenomenal knowledge is important to moral deliberation.

In chapter 4, I argued that by giving us phenomenal knowledge of another person's affective states, empathy also gives us knowledge of the intrinsic significance of those affective states for that person's well-being. When we are empathizing with a suffering person, my claim was, we will know how intrinsically bad their affective states are for them. I called this claim the Prudential Epistemic

Claim. According to this claim, empathy is an important way of finding out when people are in need and about the extent to which they are in need.

Finally, in Chapter 5, I argued for the Normative Epistemic Claim. This claim says that empathy with a suffering person gives us knowledge of the strength of our purely phenomenal reason to alleviate their suffering. The purely phenomenal reason that we have to alleviate their suffering is the reason that we have to alleviate their suffering that is given to us solely by the phenomenal quality of their suffering. If we help a suffering person for a purely phenomenal reason, we help them simply because their suffering feels awful. I also defended the Strong View, which says that our purely phenomenal reasons to alleviate suffering are stronger than we typically think they are. By combining the Strong View with the Phenomenal Epistemic Claim, I concluded that empathy has an important role to play in correcting our tendency to underestimate the strength of our purely phenomenal reasons to alleviate the suffering of other people.

To sum up, I have shown that empathy contributes more to moral deliberation than attributive knowledge. It is uniquely able to give us phenomenal knowledge about the suffering of another person. By giving us this knowledge, empathy gives us the further knowledge of how intrinsically bad that suffering is for the other person. Finally, empathy also gives us knowledge of the strength of our purely phenomenal reason to alleviate the other person's suffering. Although it is possible to acquire these latter two pieces of knowledge through other means, the Strong View implies that we typically do not. Accordingly, there is a critical role for empathy to play in remedying this deficit in our normative knowledge.

At this point, one might object to my argument by saying that I have not yet addressed the objection that empathy can lead our decision-making astray. I now use the conclusions of the previous chapter (the Strong View and the Normative Epistemic Claim) to respond to that objection.

3. Responding to the Spotlight Effect Objection

In the first chapter of this thesis, I presented an objection that has recently been made to empathy which I called the Spotlight Effect Objection (Bloom, 2016: Ch. 1; Prinz, 2011b: pp. 227-230). Roughly speaking, this objection says that empathy can push our moral decision-making off course by making us falsely believe that we should help the person that we are empathizing with, when in fact there are other people that we should be helping instead.

In my initial discussion of this objection, I illustrated it using an example taken from one of Daniel Batson's social psychology experiments (Batson, 1995). In that example, there is a ten year old girl called Sheri who has a slow-progressing terminal illness and who is on a waiting list for medical care. There are many other children on that list with similar illnesses, and each of them has been ranked according to their priority for treatment. Sheri is not at the top of the list. The worry raised by the Spotlight Effect Objection is that people who empathize with Sheri will promote her to the top of the list at the expense of those (higher priority children) who were previously above her.¹⁴⁵ In other words, the

¹⁴⁵ As I discussed in chapter 1, this worry is confirmed by the results of Batson's study. Subjects who were encouraged to empathize with Sheri (rather than 'staying objective' in response to her story) were much more likely to prioritize her for medical treatment.

objection says, the empathizer's attention is like a spotlight that shines only on Sheri and thus fails to illuminate the needs of the other children.

In chapter one, I distinguished between two versions that the Spotlight Effect Objection might take. In its 'weak' form, the objection says only that empathy can make us falsely believe that helping the person that we are empathizing with is what we have most reason to do. In its 'strong' form, the objection starts with this complaint, but then adds to it by saying that empathy leads us astray in this way by making us overstate the strength of the reason that we have to help the person that we are empathizing with. When applied to the case of Sheri, the weak version of the Spotlight Effect Objection says only that that empathizing with Sheri will make a person judge that they should put Sheri at the top of the treatment list. The strong version of the Spotlight Objection agrees that empathizing with Sheri will have this effect and adds that it has this affect because it makes the empathizer overestimate the strength of the reason that they have to help Sheri.

Having restated the Spotlight Effect Objection, and distinguished between the two forms it can take, I will now show how I think the supporter of empathy should respond to it. As a starting point, recall that the Normative Epistemic Claim (defended in the previous chapter) says that empathizing with a suffering person will give the empathizer knowledge of the strength of their purely phenomenal reason to help that person. Accordingly, the Normative Epistemic Claim implies that the strong version of the Spotlight Effect Objection is false.¹⁴⁶ If empathizing

¹⁴⁶ Technically, the strong version of the Spotlight Effect Objection could still be true if empathy for a suffering person made the empathizer overstate another kind of reason that they had to help the suffering person (i.e. a reason that was not their purely phenomenal reason to help them). I ignore this possibility here.

with Sheri prompts the empathizer to make an upward revision in their judgement about the strength of their purely phenomenal reason to help Sheri, the Normative Epistemic Claim says that this revision is appropriate. It is appropriate because the reason must have been stronger than the empathizer previously thought it was. By empathizing with Sheri, the empathizer comes to know how awful her suffering is, and thus comes to know how important it is to help her.

What about the weak version of the Spotlight Effect Objection? As I noted in the previous chapter, the Normative Epistemic Claim is consistent with that version of the objection. It is possible for an empathizer to know the strength of the purely phenomenal reason that they have to help the person that they are empathizing with and at the same time to falsely believe that helping that person is what they have most reason to do. This scenario can arise, for example, if the empathizer is underestimating the strength of the reasons that they have to help the people that they are not empathizing with.

Indeed, by recalling the Strong View, we can see why such an error might tend to arise. According to the Strong View, our purely phenomenal reasons to alleviate suffering are stronger than conventional morality would have us believe they are. In other words, that view says that we tend systematically to underestimate the strength of our purely phenomenal reasons to alleviate the suffering of other people. This explains why empathizing with a suffering person leads to an upward revision in our judgement about the strength of the reason that we have to help them. It can also explain how empathy can lead to errors in our judgements about whom to help first when multiple people are in need.

To illustrate this point, let us return to the case of Sheri. The Strong View implies that in this case, we have strong purely phenomenal reasons to alleviate the suffering of every child on the treatment list. Those reasons will be stronger in relation to the children who are higher up on the list because their suffering is worse.¹⁴⁷ Part of what is meant in saying these reasons are strong is that they are stronger than we typically think they are. The implication is that, prior to empathizing with Sheri, we are likely to be making serious normative errors. To be specific, we will be underestimating the strength of our purely phenomenal reasons to alleviate the suffering that each child is experiencing. This is true even though we also know which child we have most reason to treat first.

According to the Normative Epistemic Claim, when we empathize with Sheri, we come to know the full strength of our purely phenomenal reason to help her. As we upgrade our assessment of how important it is to help Sheri, we may be led to believe falsely that we should prioritize helping her instead of helping the other children. Indeed, we are likely to believe this if we have not also upgraded our assessment of how important it is to help those other children. As a default, the Strong Claim suggests, we will tend to underestimate this importance. Even if empathy can lead us to this error in judgement about whom we should treat first, however, it only does this by providing us with new normative knowledge.

Nevertheless, one might reply, a little bit of knowledge is a dangerous thing. A critic may push the point that empathy can still lead us astray. I accept this point, but maintain that empathy need not lead us astray if we learn how to

¹⁴⁷ I am making a simplification here. In reality, the severity of a child's suffering is only one of the factors which determines their ranking on a treatment list. Other factors might include: how long the child has been waiting for treatment, how old the child is, and how amenable their illness is to treatment. If these factors are relevant, the empathizer should also take them into account when deliberating about whom they should prioritize for treatment.

use it appropriately. When one has empathized with a suffering person, I suggest, one should make sure that one takes a step back from the insights acquired through that empathy so that one can consider one's other reasons for action. One should ask: who else is there to help? Moreover, once the empathizer has learned that they were previously underestimating the strength of their reason to help the person that they have empathized with, they should also wonder: am I underestimating the strength of my reasons to alleviate anyone else's suffering? In the Sheri case, the empathizer could thus reflect on the fact that the other children on the treatment list are probably also feeling painful affective states that are similar to those felt by Sheri. Even if the empathizer does not also empathize with those children, they can still upgrade their assessment of how important it is to help them. In this way, the empathizer can maintain a correct judgement about which child they have most reason to help. The fact that many of the subjects in Batson's experiment did not use this method to integrate empathy into their moral deliberation does not mean that it cannot be done.

At this stage, a critic may acknowledge my point that a 'skilful' use of empathy would not lead the empathizer's moral deliberations astray, but persist in asking what the point of empathy is. They may say that, prior to empathizing with Sheri, we already knew which child to treat first, and so we did not need to empathize with Sheri. In response to this, I make two points. First, the Sheri case is unusual in so far as it assumes that we already know whom we have most reason to help. Most cases that we encounter through the course of our lives will not be like this. When deliberating about whether to help a friend in need, or how much of one's money to devote to charity, there are no clear available answers. By teaching us about the strength of our purely phenomenal reasons to help other

people, empathy can help us to make these decisions. Secondly, even in the Sheri case, it could be that empathy with Sheri leads the empathizer to form an improved judgement about how many resources should be allocated to treating children like Sheri. By realizing the significance of Sheri's suffering, and thus of the suffering of the other children on the list, the empathizer could come to believe correctly that funding should be increased for the treatment program.¹⁴⁸

Generally speaking, my argument is that empathy is a vital input into moral deliberation in so far as it corrects our tendency to underestimate the strength of the purely phenomenal reasons that we have to alleviate the suffering of other people. This tendency, I have suggested, is present even in altruistically minded people. Empathy has an even bigger role to play in correcting the views of people who are inclined towards egoism and in correcting the views of people who discriminate against particular racial and social groups.

I should clarify that I not claiming that empathy is the only way of coming to know the strength of our purely phenomenal reasons to alleviate suffering. Perhaps, for example, we can also acquire this knowledge through reasoning.¹⁴⁹ My claim is, however, that empathy is one way of knowing about these reasons, and that this is an important thing to be said in favour of empathy.

Moreover, I now go on to argue, empathy is not only vital in informing our deliberations about when to alleviate the suffering of other people. It can also

¹⁴⁸ That said, it is important to note that these resources will need to be taken away from somewhere else. Just as the empathizer should not blindly prioritize the needs of the person that they are empathizing with, they also should not blindly prioritize the needs of the group to which that person belongs. If there are pressing reasons to use the resources elsewhere, then the empathizer should also attend to those.

¹⁴⁹ I have in mind here acquaintances who were convinced to dedicate their lives to alleviating suffering after reading the arguments of Peter Singer.

inform both our deliberations about when to avoid causing harm to other people and our deliberations about when to bring them joy.

4. Empathy and Harm Inhibition

Empathy is normally framed as a response to what another person is presently feeling. Indeed, I have so far focused on cases of empathizing with the present suffering of other people. However, we can also empathize with what another person would feel in a hypothetical future scenario that one of our actions can bring about.¹⁵⁰ Thus, we can empathize with a person with respect to the suffering that we could cause them. For example, a parent could empathize with their child with respect to the distress that the child would feel if the parent reprimanded them. Similarly, a politician considering invading another country could empathize with the various forms of suffering that a typical citizen of that country might endure if such an invasion took place.

What I wish to suggest in this section is that empathy can offer us an important input into our moral deliberations about whether to perform actions that harm other people. To be specific, I suggest that empathy gives us knowledge of the strength of the purely phenomenal reasons that we have to avoid causing suffering in other people. When I introduced the notion of a purely phenomenal reason in the previous chapter, I said that it was a type of reason that we had to alleviate the suffering. I am now adding that it can also be a type of reason that we have to avoid causing suffering. It is a type of reason that is given to us solely

¹⁵⁰ We could, for example, generate such empathy through perspective-taking (by imagining how the other person would feel in the situation created by our action).

by the phenomenal quality of the suffering in question. If we refrain from harming a person for a purely phenomenal reason, then we can explain our decision by saying that the other person's suffering would have felt awful.

Indeed, just as our purely phenomenal reasons to alleviate suffering are strong, so too are our purely phenomenal reasons to avoid inflicting suffering on others. Similarly, I claim that empathy is also a way of learning about these reasons. Accordingly, we could formulate a revised version of the Normative Epistemic Claim that would say that when we empathize with another person with respect to their possible future suffering, we know the strength of the purely phenomenal reason that we have to avoid causing that suffering.

For the sake of clarity, it is important to note that I am not suggesting here that we should never knowingly inflict suffering on another person. There may be cases in which, all things considered, the action that we have most reason to do involves harming another person. To return to the examples offered above: it could be argued that reprimanding a child will help them to learn, or that invading a country is necessary to prevent the government of that country from perpetrating human rights abuses. Consequently, it would be a mistake for an empathizer to jump to the conclusion that they should avoid harming the person that they are empathizing with, without first considering the wider range of reasons that they face. Instead, as with deliberations about helping other people, the empathizer should take the normative knowledge that they acquire from empathy and integrate it into their wider process of moral deliberation. Nevertheless, empathy is important to our deliberations about harming others

because it ensures that we do not understate the strength of the purely phenomenal reasons that we have to avoid doing so.¹⁵¹

To conclude this section, let us consider a case of a person who routinely and deliberately harms another person for no good reason. Specifically, and following Williams (1995), let us consider the case of a husband who cruelly mistreats his wife. In this example, even having attributive knowledge of his wife's suffering is not enough to deter this husband from mistreating her. In other words, even though he is able to identify the different forms of physical and emotional pain that she feels, this is not enough to stop him from causing this pain. Nevertheless, I suggest that if the husband could be brought to empathize with his wife, this empathy would induce a change in the husband's motivations by giving him new knowledge.¹⁵² If the husband empathized with his wife, he would come to know what her pain felt like. Knowing this, the husband would then come to know the strength of the purely phenomenal reason that he had to stop mistreating her.¹⁵³

5. Empathy with Joy

So far in this thesis, I have focused on using empathy as a response to the suffering of other people. Such empathy, I have argued, is important. There is a

¹⁵¹ If I am right in my suggestion that empathy has this role to play in inhibiting harm, one might hope to find evidence in social psychology that empathy does indeed inhibit harm. Such evidence, unfortunately, is not forthcoming because studies that explore this issue tend to give a definition of empathy that is different from mine. For more detail, see Maibom (2012).

¹⁵² Accordingly, someone who adopts Internalism about Reasons (see the previous chapter) can maintain that the husband has a reason to stop mistreating his wife, even if he does not actually empathize with her.

¹⁵³ As discussed in the previous chapter, it is possible, but unlikely, that the husband may empathize with his wife's pain and yet still fail to be motivated to stop hurting her.

lot of suffering and we have strong purely phenomenal reasons to alleviate that suffering. Empathizing with people who are suffering is a way of learning about these reasons.

We can, however, also empathize with the joy of other people. By this, I mean that we can empathize with the pleasant affective states (moods, emotions and physical sensations) that other people feel.¹⁵⁴ Examples of pleasant affective states include: excitement, gratitude, relaxation, delight, happiness, warmth, and hopefulness. These states might arise, for example, when a person meets their needs for love, sexual intimacy, belonging or meaningful work. We can empathize with the joy of other people in the same way that we empathize with their suffering. That is, we can empathize with the joy of another person by feeling a joy that is similar to theirs, introspecting our 'similar joy' and having a conscious and justified belief that they are experiencing something 'like this'.

The Strong Phenomenal Epistemic Claim implies that if we want to have phenomenal knowledge about the joy of another person, we must empathize with that person. The general version of the Prudential Epistemic Claim says that if we empathize with another person's joy, we will know the intrinsic goodness of that joy for them. In other words, that claim says that the empathizer will know how much the other person's joy, in itself, contributes to that person's well-being.¹⁵⁵

Recalling the argument of the previous chapter, I suggest that in addition to having strong purely phenomenal reasons to alleviate suffering, we also have

¹⁵⁴ To mirror my definition of suffering, I define joy to be the combination of a pleasant affective state with a liking of that state. Accordingly, I should say here that when we empathize with the joy of another person, we are empathizing with them with respect to such a combination.

¹⁵⁵ As I clarified in chapter 4, I do not mean to endorse Hedonism here. My view is that joy is one type of thing that is intrinsically good for a person, not that it is the only type of thing that is intrinsically good for a person.

strong purely phenomenal reasons to promote joy. We can know that we have such reasons by impersonally introspecting our experiences of joy. That is, if we introspect an experience of joy and relate to it simply as an experience of joy (and not as our experience of joy), then we will know that we have a reason to prolong and promote that feeling simply because of what it feels like. We will know too that we have a reason of the same strength to promote a similar feeling in any other person.

Continuing with this approach, I suggest that when we empathize with another person who is feeling joy (by feeling a joy that is similar to that which they are feeling and introspecting it), we will know the strength of our purely phenomenal reason to help that person continue to feel such joy. This idea is a new version of the Normative Epistemic Claim, adjusted to concern what we learn when we empathize with joy (rather than concerning what we learn when we empathize with suffering). By way of illustration, this claim says that if I empathize with the satisfaction and excitement that a theatre workshop is bringing its participants, I will know the strength of my purely phenomenal reasons to keep running the workshop.

In the previous section, I suggested that we can empathize with the suffering that we can cause in other people. Similarly, we can also empathize with the joy that we can cause in other people. Doing this, I suggest, is a way of learning about the strength of our purely phenomenal reasons to help other people achieve such joy. For example, if I was not already running a theatre workshop, but expected that such a workshop would give joy to its participants, then empathizing with that prospective joy would give me knowledge of my purely phenomenal reasons to start running such a workshop. By the same logic,

empathizing with how a friend would feel if I gave him a thoughtful present gives me knowledge of the strength of my purely phenomenal reason to do so.

As a final point in this section, it is important to note that where a person is presently suffering, then we might hope that by helping this person we would not only alleviate their suffering, but also bring them joy. For example, if a friend is feeling depressed, I would hope that my supportive actions would not only help alleviate his depression, but also help him to feel excited and inspired. If I empathize my friend only with respect to his present suffering, I will only gather half of the story about my purely phenomenal reasons to help him. To be more specific, I will only learn about my purely phenomenal reasons to alleviate his suffering. Yet I also have purely phenomenal reasons to promote his joy. To recognize both of these sets of reasons, I must empathize with my friend with respect to his present suffering and with respect to the joy that I could help him to feel.

6. Conclusion

In this thesis, I have argued that empathy is an important input into moral deliberation. Empathy teaches us about our purely phenomenal reasons to alleviate the present suffering of other people, and about our purely phenomenal reasons to avoid causing the suffering of other people. It also teaches us about our purely phenomenal reasons to help other people to feel joy. These are all reasons that, without the use of empathy, we tend to understate.

Before closing, it is worth noting an additional benefit to empathizing with the joy of other people. Whereas empathizing with the suffering of other people is

painful and draining (as asserted by the Emotional Cost Objection), empathizing with the joy of other people is pleasant and uplifting. This is a point that has been acknowledged even by Paul Bloom, one of the strongest critics of empathy. At the end of his book *Against Empathy*, he makes the following concession:

Empathy can be an immense source of pleasure. Most obviously, we feel joy at the joy of others. I've noted elsewhere that here lies one of the joys of having children: You can have experiences that you've long become used to – eating ice cream, watching Hitchcock movies, riding a roller coaster – for the first time all over again. Empathy amplifies the pleasures of friendship and community, of sport and games, and of sex and romance. (Bloom, 2017: p. 241)

I agree with Bloom that empathizing with the joys of other people is a source of pleasure. Where I go beyond Bloom is in arguing that such empathy is also an important input into moral deliberation. Such empathy helps us to learn about the strong other-regarding reasons that we have, for example, to help children thrive, to forge healthy relationships, friendships and communities, and to create opportunities for sport and games. Moral deliberation is not just about alleviating suffering, it is also about cultivating joy. Empathy teaches us about our reasons to do both.

Bibliography

- Adelman, P. K. & Zajonc, R. B., 1989. Facial Expression and the Experience of Emotion. *Annual Review of Psychology*, Volume 40, pp. 249-280.
- Alston, W., 1968. Pleasure. In: P. Edwards, ed. *The Encyclopedia of Philosophy*. New York: Collier-Macmillan.
- Baron-Cohen, S., 2011. *Zero Degrees of Empathy: A New Theory of Human Cruelty*. London: Allen Lane.
- Batson, C. D., 1991. *The Altruism Question: Towards a Social-Psychological Answer*. Hillsdale, NJ: Erlbaum.
- Batson, C. D., 1999. *The Altruism Question: Toward a Social Psychological Answer*. Hillsdale, N.J.: Lawrence Erlbaum.
- Batson, C. D., 2011. *Altruism in Human*. New York: Oxford University Press.
- Batson, C. D. et al., 1981. Is empathic emotion a source of altruistic motivation?. *Journal of Personality and Social Psychology*, Volume 40, pp. 290-302.
- Batson, C. D., Early, S. & Salvarini, G., 1997. Perspective Taking: Imagining How Another Feels Versus Imagining How You Would Feel. *Personality & Social Personality Bulletin*, Volume 23, pp. 751-8.
- Batson, C. D., Klein, T. R., Highberger, L. & L, S. L., 1995. Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and Social Psychology*, Volume 68, pp. 1042-54.
- Battaly, H. D., 2011. Is Empathy a Virtue?. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- Bazalgette, P., 2017. *The Empathy Instinct: How to Create a More Civil Society*. London, UK: John Murray.
- Beecher, H. K., 1959. *Measurement of Subjective Responses*. New York: Oxford University Press.
- Bending, L., 2006. Approximation, Suggestion and Analogy: Translating Pain Into Language. *The Yearbook of English Studies*, 36(1), pp. 131-137.
- Berger, M. S. & Hadley, S. W., 1975. Some Effects of a Model's Performance on an Observer's Electromyographic Activity. *American Journal of Psychology*, 88(2), pp. 263-76.
- Bernhardt, B. C. & Singer, T., 2012. The Neural Basis of Empathy. *The Annual Review of Neuroscience*, Volume 35, pp. 1-23.
- Berridge, K. & Winkielman, P., 2003. What is an unconscious emotion? (The case for unconscious "liking"). *Cognition and Emotion*, 17(2), pp. 181-211.

- Biggs, S., 2007. The Phenomenal Mindreader: A Case for Phenomenal Simulation. *Philosophical Psychology*, 20(1), pp. 29-42.
- Biggs, S., 2009. Phenomenal concepts in mindreading. *Philosophical Psychology*, 22(6), pp. 647-667.
- Blair, R. J. R., 1995. A cognitive development approach to morality: Investigating the psychopath. *Cognition*, Volume 57, pp. 1-29.
- Block, N., 1990. Inverted Earth. *Philosophical Perspectives*, Volume 4, pp. 53-79.
- Block, N., 1994. Qualia. In: S. Guttenplan, ed. *Blackwell Companion to the Philosophy of Mind*. Oxford: Blackwell.
- Bloom, P., 2013. *The Baby in the Well: The case against empathy*. [Online] Available at: <http://www.newyorker.com/magazine/2013/05/20/the-baby-in-the-well> [Accessed 08 Mar 2016].
- Bloom, P., 2014. *Against Empathy*. [Online] Available at: <https://bostonreview.net/forum/paul-bloom-against-empathy> [Accessed 08 Mar 2016].
- Bloom, P., 2016. *Against Empathy*. New York, USA: HarperCollins.
- Botvinick, J. A. P. et al., 2005. Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage*, Volume Vol. 25, pp. 312-9.
- Brandt, R., 1979. *A Theory of the Good and the Right*. Oxford: Clarendon Press.
- Brink, D., 1997. Rational Egoism and the Separateness of Persons. In: J. Dancy, ed. *Reading Parfit*. Oxford: Blackwell Publishing.
- Brooks, D., 2011. *The Limits of Empathy*. [Online] Available at: http://www.nytimes.com/2011/09/30/opinion/brooks-the-limits-of-empathy.html?_r=0 [Accessed 08 Mar 2016].
- Brown, L. M., Bradley, M. & Lang, P., 2006. Affective Reactions to Pictures of Ingroup and Outgroup Members. *Biological Psychology*, Volume 71, pp. 303-11.
- Carruthers, P., 2013. Mindreading in Infancy. *Mind and Language*, pp. 141-172.
- Carse, A. L., 2005. The Moral Contours of Empathy. *Ethical Theory and Moral Practice*, 8(1-2), pp. 169-195.
- Carson, T., 2000. *Value and the Good Life*. Notre Dame: University of Notre Dame Press.
- Cassidy, J., 1979. Some Similarities Between Hume's and Spinoza's Ethical Theories. *The Journal of Value Inquiry*, pp. 187-194.

- Chalmers, D., 2003. The content and epistemology of phenomenal belief. In: Q. Smith & A. Jokic, eds. *Consciousness: New philosophical perspectives*. New York: Oxford University Press.
- Chalmers, D. J., 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Churchland, P., 1985. Reduction, Qualia and the Direct Introspection of Brain States. *Journal of Philosophy*, Volume 82, pp. 8-28.
- Cialdini, R. B. et al., 1997. Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, Volume 73, pp. 481-494.
- Clark, A. J., 2007. *Empathy in Counseling and Psychotherapy*. New York: Routledge.
- Conee, E., 1994. Phenomenal Knowledge. *Australasian Journal of Philosophy*, Volume 72, pp. 136-150.
- Coplan, A., 2011. Understanding Empathy. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- Coplan, A. & Goldie, P. eds., 2011. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- Crisp, R., 2006. *Reasons and the Good*. New York: Oxford University Press.
- D'Arms, J., 2000. Empathy and Evaluative Inquiry. *Chicago Law Review*, Volume 74, pp. 1467-1500.
- D'Arms, J., 2011. Empathy, Approval and Disapproval in Moral Sentimentalism. *Southern Journal of Philosophy*, Volume 49 (s1), pp. 134-141.
- Darwall, S., 1998. Empathy, Sympathy, Care. *Philosophical Studies*, Volume 89 (2-3), pp. 261-282.
- Darwall, S., 2004. Respect and the Second-Person Standpoint. *Proceedings and Addresses of the American Philosophical Association*, Volume 82 (2), pp. 43-59.
- Darwall, S., 2006. *The Second Person Standpoint: Morality, Respect and Accountability*. Cambridge: Harvard University Press.
- Darwall, S., 2011. Being With. *Southern Journal of Philosophy*, Volume 49 (s1), pp. 4-24.
- Daudet, A., 2002. *In the Land of Pain*. London: Cape.
- De Waal, F., 2009. *The Age of Empathy: Nature's Lessons For A Kinder Society*. London: Souvenir Press.
- Decety, J. & Sommerville, J. A., 2003. Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Science*, Volume 7 (12), pp. 527-33.

- Deigh, J., 1996. *The Sources of Moral Agency: Essays on Moral Psychology*. Cambridge: Cambridge University Press.
- Denham, A., 2017. Empathy and Moral Motivation. In: *The Routledge Handbook of Philosophy of Empathy*. Abingdon, UK: Routledge, pp. 227-242.
- Dennett, D., 2002. Quining Qualia. In: D. Chalmers, ed. *Philosophy of mind: Classical and contemporary readings*. Oxford: Oxford University Press.
- Depew, D., 2005. Empathy, Psychology and Aesthetics: Reflections on a Repair Concept. *Poroi*, 4(1), pp. 99-107.
- Dimberg, U., 1982. Facial Reactions to Facial Expressions. *Journal of Psychophysiology*, 19(6), pp. 643-647.
- Dimberg, U., 1988. Facial Electromyography and the Experience of Emotion. *Journal of Psychophysiology*, Volume 56, pp. 277-288.
- Eisenberg, N., 2000. Empathy and Sympathy. In: M. Lewis & J. Haviland-Jones, eds. *Handbook of Emotion*. New York: The Guilford Press, pp. 677-91.
- Eisenberg, N. & Miller, P., 1987. Relation of Empathy to Prosocial Behaviour. *Psychological Bulletin*, Volume 101, pp. 91-119.
- Eisenberg, N. & Strayer, J., 1987. Critical Issues in the Study of Empathy. In: N. Eisenberg & J. Strayer, eds. *Empathy and Its Development*. New York: Cambridge University Press.
- Ekman, 1980. *The Face of Man: Expressions of Universal Emotions in a New Guinea Village*. London: Garland STPM Press.
- Ekman, P., 1992. Facial Expressions: New Findings, New Questions. *Psychological Science*, Volume 3, pp. 34-38.
- Ekman, P., Levenson, R. W. & Friesen, W., 1983. Autonomic Nervous System Activity Distinguishes Among Emotions. *Science, New Series*, Volume 221, pp. 1208-1210.
- Eslinger, P. J., 1998. Neurological and neuropsychological bases of empathy. *European Neurology*, pp. 193-199.
- Feldman, F., 1997. On the Intrinsic Value of Pleasures. *Ethics*, 107(3), pp. 448-466.
- Feldman, F., 2004. *Pleasure and the Good Life*. New York: Oxford University Press.
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G., 1996. Action recognition in the premotor cortex. *Brain: A Journal of Neurobiology*, 119(2), pp. 593-609.
- Gilligan, C., 1982. *In A Different Voice: Psychological Theory and Women's Development*. Cambridge: Harvard University Press.

- Goldie, P., 2002. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.
- Goldman, A., 2006. *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. New York, USA: Oxford University Press.
- Goldman, A., 2011. Two Routes to Empathy: Insights from Cognitive Neuroscience. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- Goldstein, I., 1989. Pleasure and Pain: Unconditional, Intrinsic Values. *Philosophy and Phenomenological Research*, Volume 1 (2), pp. 255-276.
- Gopnik, A. & M., W. H., 1992. Why the Child's Theory of Mind Really Is a Theory. *Mind and Language*, Volume 7 (1-2), pp. pp. 145-71.
- Gordon, M., 2005. *Roots of Empathy: Changing the World Child by Child*. New York: The Experiment Publishing.
- Gordon, R. M., 1996. Sympathy, simulation, and the impartial spectator. In: *Mind and Morals*. Cambridge, MA: MIT Press, pp. 165-180.
- Grahek, N., 1991. Objective and subjective aspects of pain. *Philosophical Psychology*, 4(2), pp. 249-266.
- Grahek, N., 2007. *Feeling Pain and Being in Pain*. MA: MIT Press.
- Green, M. S., 2007. *Self-Expression*. Oxford: Oxford University Press.
- Griffin, J. H., 1961/2009. *Black Like Me*. London: Souvenir Press.
- Hacker, P. M. S., 2002. Is there anything that it is like to be a bat?. *Philosophy*, Volume 77, pp. 154-172.
- Halpern, J., 2001. *From Detached Concern To Empathy*. New York, USA: Oxford University Press.
- Hare, R. M., 1972. Pain and Evil. *Proceedings of the Aristotelian Society, Supplementary Volumes*, Volume 38, pp. 91-124.
- Hare, R. M., 1981. *Moral thinking: its levels, method, and point*. New York: Oxford University Press.
- Harris, P., 1992. From simulation to folk-psychology: The case for development. *Mind and Language*, Volume 7, pp. 120-144.
- Heal, J., 1995. How to think about thinking. In: *Mental Simulation*. Oxford, UK: Blackwell, pp. 33-52.
- Heathwood, C., 2007. The Reduction of Sensory Pleasure To Desire. *Philosophical Studies*, Volume 133, pp. 23-44.
- Hilton, L. et al., 2017. Mindfulness Meditation for Chronic Pain: Systematic Review and Meta-analysis. *Annals of Behavioural Medicine*, 51(2), pp. 199-213.

- Hoffman, M. L., 1978. Empathy, its development and prosocial implications. *Nebraska Symposium on Motivation*.
- Hoffman, M. L., 1991. Is Empathy Altruistic?. *Psychological Inquiry*, Volume 2, pp. 131-133.
- Hoffman, M. L., 2000. *Empathy and Moral Development*. Cambridge: Cambridge University Press.
- Hoffman, M. L., 2002. Empathy and Vicarious Traumatization in Clinicians. *Unpublished Manuscript*.
- Hoffman, M. L., 2011. Empathy, Justice, and the Law. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- Hoffman, M. L., 2013. Empathy, Justice and Social Change. In: H. Maibom, ed. *Empathy and Morality*. New York: Oxford University Press.
- Horgan, T., 1984. Jackson on Physical Information and Qualia. *Philosophical Quarterly*, Volume 32, pp. 127-132.
- Howe, D., 2013. *Empathy: What it is and why it matters*. Basingstoke, UK: Palgrave Macmillan.
- Human (Extended Version)*. 2015. [Film] Directed by Yann Arthus-Bertrand. s.l.: Humankind Production.
- Hume, D., 1739/1978. *A Treatise of Human Nature*. Oxford: Oxford University Press.
- Iacoboni, M., 2008. *Mirroring People: The New Science of How We Connect With Others*. New York: Farrar, Straus and Giroux.
- Ickes, W., 1993. Empathic Accuracy. *Journal of Personality*, Volume 61, pp. 587-610.
- Ilyes, I., 2017. Empathy in Hume and Smith. In: H. Maibom, ed. *The Routledge Handbook of Empathy*. Abingdon: Routledge.
- Jabbi, M., Swart, M. & Keysers, C., 2007. Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage*, Volume 34, pp. 1744-53.
- Jackson, F., 1982. Epiphenomenal Qualia. *The Philosophical Quarterly*, 32(127), pp. 127-136.
- Jackson, F., 1986. What Mary didn't know. *The Journal of Philosophy*, Volume Vol. 82, pp. 291-295.
- Jackson, P., Meltzoff, A. & Decety, J., 2005. How do we perceive the pain of others: A window into the neural processes involved in empathy. *NeuroImage*, Volume 24, pp. 771-779.

- Kagan, S., 1992. The Limits of Well-Being. In: E. F. Paul, F. D. Miller Jr & J. Paul, eds. *The Good Life and the Human Good*. Cambridge: Cambridge University Press, pp. 169-189.
- Kahane, G., 2009. Pain, Dislike and Experience. *Utilitas*, 21(3), pp. 327-336.
- Kahane, G., 2010. Feeling pain for the very first time: The normative knowledge argument. *Philosophy and Phenomenological Research*, Volume 80, pp. 20-49.
- Kahane, G., 2015. Pain, experience, and well-being. In: G. Fletcher, ed. *The Routledge Handbook of Philosophy of Well-Being*. Abingdon: Routledge.
- Kahneman, D., Fredrickson, B. L., Schreiber, C. A. & Redelmeier, D. A., 1993. When More Pain Is Preferred to Less: Adding a Better End. *Psychological Science*, 4(6), pp. 401-405.
- Kaplan, E. A., 2011. Empathy and Trauma Culture: Imagining Catastrophe. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. New York: Oxford University Press.
- Kauppinen, A., 2014. Empathy, Emotion Regulation, and Moral Judgement. In: H. Maibom, ed. *Empathy and Morality*. Oxford: Oxford University Press.
- Keysers, K. et al., 2004. A Touching Sight: SII/PV Activation during the Observation and Experience of Touch. *Neuron*, 42(2), pp. 335-346.
- Kohut, H., 1981. Introspection, Empathy, and the Semicircle of Mental Health. In: P. H. Orstein, ed. *The Search for the Self: Selected Writings of Heinz Kohut: 1978-1981*. Madison, CT: International Universities Press, pp. 537-67.
- Krebs, D. L., 1975. Empathy and Altruism. *Journal of Personality and Social Psychology*, pp. 1134-1146.
- Krznaric, R., 2014. *Empathy: A Handbook for Revolution*. Croydon: Random House.
- Lewis, D., 1988. What Experience Teaches. *Proceedings of the Russellian Society*, Volume 13, pp. 29-57.
- Lhermitte, F., Pillon, B. & Serdaru, M. D., 1986. Human Autonomy and the Frontal Lobes. Part I: Imitation and Utilization Behaviour. A Neurophysiological Study of 75 Patients. *Annals of Neurology*, 19(4), pp. 326-334.
- Lipps, T., 1903/1931. Empathy, Inward Imitation, and Sense Feelings. In: E. F. Carritt, ed. *Philosophies of Beauty: From Socrates to Robert Bridges being the Sources of Aesthetic Theory*. Oxford: Clarendon Press, pp. 252-6.
- Lishner, D. et al., 2012. Evaluating the relation between affective psychopathy and empathy: Two preliminary studies. *International Journal of Offender Therapy and Comparative Criminology*, Volume 56, pp. 1161-1181.
- Loar, B., 1997. Phenomenal states. In: N. Block, O. Flanagan & G. Guzeldere, eds. *The nature of consciousness*. Cambridge, MA: MIT Press.

Ludlow, P., Nagasawa, Y. & Stoljar, D. eds., 2004. *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. Cambridge, MA: MIT Press.

Maibom, H., 2014. Introduction: (Almost) Everything You Ever Wanted to Know about Empathy. In: H. Maibom, ed. *Empathy and Morality*. New York: Oxford University Press.

Maibom, H. L., 2007. The Presence of Others. *Philosophical Studies*, pp. 161-190.

Maibom, H. L., 2012. The many faces of empathy and their relation to prosocial action and aggression inhibition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(2), pp. 253-263.

Maibom, H. L., 2017. Introduction. In: *The Routledge Handbook of Philosophy of Empathy*. Abingdon: Routledge.

Masten, C. L., Morelli, S. A. & Eisenberger, N. I., 2011. An fMRI investigation of empathy for 'social pain' and subsequent prosocial behaviour. *Neuroimage*, Volume 55, pp. 381-88.

Masto, M., 2015. Empathy and Its Role in Morality. *The Southern Journal of Philosophy*, 53(1), pp. 74-96.

Matravers, D., 2011. Empathy as a Route to Knowledge. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.

McIntosh et al., D., 1994. Socially Induced Affect. In: D. Druckman & R. A. Bjork, eds. *Learning, Remembering, Believing: Enhanced Human Performance*. Washington: National Academy Press, pp. 251-277.

McIntosh, D., 1994. Socially Induced Affect. In: D. Druckman & B. R. A, eds. *Learning, Remembering, Believing: Enhancing Human Performance*. Washington: National Academies Press.

Meltzoff, A. N. & Moore, K. M., 1983. Newborn Infants Imitate Adult Facial Gestures. *Child Development*, 54(3), pp. 702-709.

Mendola, J., 1990. Objective Value and Subjective States. *Philosophy and Phenomenological Research*, Volume 50 (4), pp. 695-713.

Mendola, J., 2006. Intuitive Hedonism. *Philosophical Studies*, Volume 128, pp. 441-477.

Miller, C. B., 2013. *Moral Character: An Empirical Approach*. Oxford: Oxford University Press.

Morton, A., 2011. Empathy for the Devil. In: C. A & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.

- Moser, C., 1972. *Columbia Essays on Modern Writers*. New York: Columbia University Press.
- Nagel, 1974. What is it like to be a bat?. *The philosophical review*, 83(4), pp. 435-450.
- Nelkin, N., 1994. Reconsidering Pain. *Philosophical Psychology*, 7(3), pp. 325-343.
- Nemirow, L., 1990. Physicalism and the Cognitive Role of Acquaintance. In: W. G. Lycan, ed. *Mind and Cognition*. Oxford: Blackwell.
- Nichols, S., 2001. Mindreading and the cognitive architecture underlying altruistic motivation. *Mind & Language*, Volume 16, pp. 425-55.
- Nichols, S., 2004. *Sentiment Rules: On the Natural Foundations of Moral Judgement*. New York: Oxford University Press.
- Nillson, P., 2003. *Empathy and Emotions: On the Notion of Empathy as Emotional Sharing (Doctoral Dissertation)*. Umea University: Department of Philosophy and Linguistics.
- Noddings, N., 1984. *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley, CA: University of California Press.
- Nordby, K., 2007. What is this thing you call color? Can a totally color blind person know about color?. In: T. Alter & S. Walter, eds. *Phenomenal Concepts and Phenomenal Knowledge New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.
- Nussbaum, M. C., 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Orwell, G., 1962. *The Road to Wigan Pier*. Harmondsworth: Penguin.
- Oxley, J., 2012. *The Moral Dimensions of Empathy: Limits and Applications in Ethical Theory and Practice*. London: Palgrave MacMillan.
- Papineau, D., 2002. *Thinking about Consciousness*. Oxford: Oxford University Press.
- Parfit, D., 1997. Reasons and Motivation. *Proceedings of the Aristotelean Society, Supplementary Volumes*, Volume 71, pp. 99-146.
- Ploner, M., Freund, H.-J. & Schnitzler, A., 1999. Pain Affect Without Pain Sensation in a Patient with a Postcentral Lesion. *Pain*, Volume 81, pp. 211-214.
- Prehn-Kristensen, A. et al., 2009. Induction of Empathy by the Smell of Anxiety. *PLOS One*, Volume Vol. 4: e5987.
- Prinz, J., 2011a. Is Empathy Necessary for Morality?. In: A. Coplan & P. Goldie, eds. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.

- Prinz, J., 2011b. Against Empathy. *The Southern Journal of Philosophy*, Volume 49 (s1), pp. 214-233.
- Rachels, S., 2000. Is Unpleasantness Intrinsic to Unpleasant Experiences?. *Philosophical Studies*, Volume 99, pp. 187-210.
- Ravenscroft, I., 1998. What is it like to be someone else? Simulation and Empathy. *Ratio*, Volume 11, pp. 170-185.
- Ravenscroft, I., 2017. Empathy and "knowing what it's like". In: H. Maibom, ed. *The Routledge Handbook of the Philosophy of Empathy*. Abingdon: Routledge.
- Rifkin, J., 2009. *The Empathic Civilization: The Race to Global Consciousness in a World in Crisis*. London: Penguin Group.
- Rogers, C., 1980. *A Way of Being*. New York: Houghton Mifflin.
- Rosenberg, M., 2003. *Nonviolent Communication: A Language of Life*. Encinitas, CA: PuddleDancer Press.
- Scarry, E., 1985. *The Body in Pain*. Oxford: Oxford University Press.
- Scheler, M., 1970/1979. *The Nature of Sympathy*. London: Routledge & Kegan Paul.
- Searle, J. R., 2000. Consciousness. *Annual Review of Neuroscience*, Volume 23, pp. 557-578.
- Shoemaker, S., 1982. The Inverted Spectrum. *Journal of Philosophy*, Volume 79, pp. 357-381.
- Sidgwick, H., 1981. *The Methods of Ethics*. Indianapolis: Hackett.
- Simmons, A., 2014. In Defense of the Moral Significance of Empathy. *Journal of Ethical Theory and Moral Practice*, pp. 97-111.
- Singer, T. a. K. O. M., 2014. Empathy and Compassion. *Current Biology*, pp. R875-878.
- Singer, T. & Lamm, C., 2009. The Social Neuroscience of Empathy. *The Year in Cognitive Neuroscience: Annals of the New York Academy of Sciences*, Volume 1156, pp. 81-96.
- Singer, T. et al., 2004. Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, Volume 303, pp. 1157-1162.
- Sinhababu, N., 2012. The Epistemic Argument for Hedonism. *Unpublished manuscript*.
- Slote, M., 2007. *The Ethics of Care and Empathy*. Abingdon: Routledge.
- Slote, M., 2010. *Moral Sentimentalism*. Oxford: Oxford University Press.
- Smith, A., 1759/1975. *The Theory of Moral Sentiments*. Oxford: Clarendon.
- Smith, J., 2017. What is empathy for?. *Synthese*, 194(3), pp. 709-722.

- Smuts, A., 2011. The feels good theory of pleasure. *Philosophical Studies*, 155(2), pp. 241-265.
- Sober, E. & Wilson, D. S., 1998. *Unto Others*. Cambridge, MA: Harvard University Press.
- Song, Y., 2015. How to Be a Proponent of Empathy. *Ethical Theory and Moral Practice*, 18(3), pp. 437-451.
- Spinoza, B., 1677/1994. Ethics. In: E. Curley, ed. *A Spinoza Reader: The Ethics and Other Works*. Princeton: Princeton University Press.
- Spinrad, T. L. & Eisenberg, N., 2013. Empathy and Morality - A Developmental Psychology Perspective. In: *Empathy and Morality*. New York: Oxford University Press, pp. 60-71.
- Stein, E., 1916/1989. *On the Problem of Empathy (The Collected Works of Edith Stein, Volume 3)*. 3rd Revised Edition ed. Washington, D.C. : ICS Publications.
- Stotland, E., 1969. Exploratory Investigations of Empathy. *Advances in Experimental Social Psychology*, Volume 4, pp. 272-314.
- Stowe, H., 2009. Autobiographical Letter to Eliza Cabot Fallon. In: *Stowe in Her Own Time*. Iowa City: University of Iowa Press, pp. 62-9.
- Stueber, K., 2006. *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge, Mass. : MIT Press.
- Stürmer, S., Snyder, M. & Omoto, A., 2005. Prosocial Emotions and Helping: The Moderating Role of Group Membership. *Journal of Personality and Social Psychology*, Volume 88, pp. 532-546.
- Sumner, L. W., 1999. *Welfare, Happiness and Ethics*. Oxford: Oxford University Press.
- Tannsjo, T., 2007. Narrow Hedonism. *Journal of Happiness Studies*, pp. 79-98.
- Titchener, E. B., 1909. *Lectures on the Experimental Psychology of Thought-Processes*. New York: Macmillan.
- Toi, M. & Batson, C. D., 1982. More evidence that empathy is a source of altruistic motivation. *Journal of Personality and Social Psychology*, Volume 43, pp. 282-292.
- Trout, J. D., 2009. *Why Empathy Matters: The Science and Psychology of Better Judgment*. New York, USA: Penguin Group.
- Tye, M., 1986. The Subjective Qualities of Experience. *Mind*, 95(377), pp. 1-17.
- Tye, M., 1995. *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- Tye, M., 2000. Knowing What It Is Like: The Ability Hypothesis and the Knowledge Argument. In: G. Preyer, ed. *Reality and Humean Supervenience*:

- Essays on the Philosophy of David Lewis*. Lanham(Maryland): Rowman & Littlefield.
- Tye, M., 2012. Knowing What It Is Like. In: J. Bengson & M. A. Moffett, eds. *Knowing How: Essays on Knowledge, Mind, and Action*. Oxford: Oxford University Press.
- Underwood, B. & Moore, B., 1982. Perspective-taking and Altruism. *Psychological Bulletin*, Volume 91, pp. 43-73.
- Van Bowen, L. & G, L., 2003. Social Projection of Transient Drive States. *Personality and Social Psychology Bulletin*, Volume 29 (9), pp. 1159-68.
- Vaughan, K. B. & Lanzetta, J. T., 1981. The effect of modification of expressive displays on vicarious emotional arousal. *Journal of Experimental Social Psychology*, pp. 16-30.
- Vignemont, F. d. & Singer, T., 2006. The Empathetic Brain: How, When and Why?. *Trends in Cognitive Sciences*, Volume 10, pp. 435-441.
- Vischer, R., 1873. On the Optical Sense of Form: A Contribution to Aesthetics. In: H. F. Mallgrave & E. Ikonomou, eds. *Empathy, Form, and Space*. Santa Monica, CA: The Getty Center for the History of Art and the Humanities, pp. 89-123.
- Wallraf, G., 1988. *Lowest of the Low*. London: Meuthen.
- Weil, S., 1951/2009. *Waiting on God*. New York: HarperCollins.
- Wicker, B. et al., 2003. Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*, Volume 40, pp. 655-64.
- Williams, B., 1979. Internal and External Reasons. In: R. Harrison, ed. *Rational Action*. Cambridge: Cambridge University Press.
- Williams, B., 1995. Internal reasons and the obscurity of blame. In: *Making sense of humanity*. Cambridge, MA: Cambridge University Press.
- Williams, M. & Penman, D., 2011. *Mindfulness: A Practical Guide to Finding Peace in a Frantic World*. London: Piatkus.
- Zahavi, D., 2014. *Self and Other: Exploring Subjectivity, Empathy and Shame*. Oxford: Oxford University Press.