

Empathy and Moral Judgment

Antti Kauppinen

Revised draft, April 28, 2016

For *Routledge Handbook of the Philosophy of Empathy*, ed. Heidi Maibom.

At the beginning of September 2015, shocking images of a drowned Syrian refugee boy on a Turkish beach aroused widespread criticism of European policy. Although the crisis had started much earlier, many people apparently only then formed the belief that it is the moral obligation of rich Europe to take care of people in desperate need, and demanded politicians to act. Why? Speaking for myself, as a parent of a boy of similar age, I felt sadness and anger at those responsible for forcing parents to take such risks – I couldn't help thinking that this could have happened to my own son, had I not had the luck to live in a stable and peaceful country. The striking picture resonated emotionally with me, as it did with many others who had hitherto paid little attention to the refugee problem, in spite of knowing that large numbers of people were risking their lives to escape war. The best explanation for this reaction is likely to be the capacity and tendency of human beings to take on the feelings they attribute to other people, when they come to be vividly aware of the situation of individual others they can identify with. Evidently, such empathic feelings sometimes causally influence the moral judgments that people make.

This modest claim about causal influence is hardly controversial. But some philosophers have made stronger claims for empathy, maintaining that it is *necessary* for or even constitutive of moral judgment, or that it is part of the best explanation of why we endorse pro-social moral norms and distinguish them from conventional norms. Some have also argued that empathy is needed for making *good* moral judgments, while others claim it's often morally problematic, because it is biased and insensitive to numbers, among other

things. This chapter examines arguments for and against such claims, focusing largely on the contemporary debate, though I will discuss the historical views that have directly influenced it.

1. Empathy Causation

In keeping with established terminology, I will distinguish between cognitive and affective empathy, where the former is roughly a matter of imaginatively taking another's perspective, and the latter involves roughly coming to feel as the other does, because one takes the other to feel that way. (I will leave the details of these processes for other chapters.) Here's a simple form of the hypothesis that empathy is causally necessary for moral judgment:

Minimal Empathy Causation Hypothesis

Any moral judgment made by any subject B regarding a situation that elicits emotion or affect in another subject A is caused at least in part by affective empathy with A.

If Minimal Empathy Causation is true, we can't make moral judgments concerning others without first empathizing with someone. (As Jesse Prinz (2011a) points out, it is silent on judgments concerning oneself.) Why would this be? One classical argument is provided by David Hume (whose own view is nevertheless ultimately more complex, as we'll see). Hume argues, first, that what makes us approve or disapprove of something is that surveying it gives rise to a distinctive kind of pleasure or pain (T 78). In the case of moral judgment, he maintains that "'Tis only when a character is consider'd in general, without reference to our particular interest, that it causes such a feeling or sentiment, as denominates it morally good or evil." (T 79) Second, something pleases or pains us without reference to our particular interest only when we empathize with the pleasure or pain it gives rise to in others. Using "sympathy" for what is now called empathy, he summarizes: "When any quality, or

character, has a tendency to the good of mankind, we are pleas'd with it, and approve of it; because it presents the lively idea of pleasure; which idea affects us by sympathy, and is itself a kind of pleasure." (T 155) This suggests that empathy is causally necessary for (other-directed) moral judgment, since it alone enables us to have the distinctive kind of disinterested pleasure or pain on which moral approbation or disapprobation is based. This is a parsimonious hypothesis, as it allows us to explain why people make the moral judgments they do without appeal to some kind of innate moral capacity, intuition, or practical reason. (Hume and other sentimentalists separately argue against these alternative explanations, but this is not the place to discuss these arguments.)

Nevertheless, Minimal Empathy Causation faces such serious challenges that it has few if any defenders. Leaving aside issues that arise for sentimentalist explanations in general, Hume himself observed that our empathy can vary without variation in our moral judgment. For example, he noted that like sense perception, our natural empathy is influenced by the position of the object relative to us: "We sympathize more with persons contiguous to us, than with persons remote from us: With our acquaintance, than with strangers: With our countrymen, than with foreigners." (T 156) Yet our approval doesn't (always) vary accordingly. Hume's explanation of this was that we learn to *regulate* our empathy when making judgments. That is, we "correct the momentary appearances" (T 157) by adopting a "common point of view" (T 163), since otherwise our sentiments would constantly clash and uncertainty would reign. We do this "by a sympathy with those, who have any commerce with the person we consider" (T 158). So it is not enough that empathy helps us transcend our own perspective in *some* way. For Hume, our moral verdicts depend on empathizing with the feelings of those affected by an action (or the agent's character traits) regardless of their relationship to us. Call this kind of view the Regulated Empathy Causation hypothesis.

Adam Smith's account builds on Hume's, but highlights the role of *cognitive* empathy. He believes that what leads us to approve of someone's response to a situation is that we imaginatively place ourselves in their shoes and find that we would respond the same way. His account of moral judgment is complex, but roughly, he claims that we morally disapprove of someone if we imaginatively place ourselves in the shoes of both the agent and those affected, and find that we would ourselves resent the agent for the ill will her action displays (see Gordon 1995 and Kauppinen 2010). Like Hume, Smith thinks we learn to regulate our response, in his case by reference to how an impartial spectator would feel. An impartial spectator is just any normal person who doesn't favor any particular person, so that her responses are not influenced by the identity of the agent or the patient of the action, and who doesn't think of herself as more important than others. When I approach a situation as an impartial spectator, I feel just the same way about an insult to a stranger as I do about an insult to a friend, and I take it that any ordinary person who treats others as equals would feel the same way. It is the sense that *any* normal, decent person who doesn't take sides would feel in a certain way that lends a distinctively moral force and quality to our sentiment.

So, Hume and Smith agree that making moral judgments on the basis of empathy involves counteracting some of our natural tendencies. In contemporary terms, it demands a form of *emotion regulation*: we need to both *up-regulate* our empathic reaction on behalf of strangers and *down-regulate* our empathic reaction on behalf of those close to us (see Kauppinen 2014 for discussion with reference to empirical psychology). However, while this kind of view explains why natural empathy and our moral judgments diverge, it is hardly credible as a hypothesis about the causal history of *each and every* moral judgment. The problem is that it is cognitively quite demanding to place ourselves impartially in another's position before judging, and we certainly seem capable of making judgments without doing so.

Even if we leave the above issues aside, it is questionable whether empathizing could account for the content of all our moral judgments. As Jesse Prinz has emphasized, there are many situations in which we make moral judgments, but there is no possibility of empathic affective reaction. For example, there seem to be victimless crimes, which we disapprove of even though there is by definition no one to empathize with. Many people disapprove of masturbating with an already dead chicken, for example (Haidt, Koller, and Diaz 1996). In yet other cases, there are too many victims to empathize with (think of the Great Famine), or it is indeterminate who suffers from the bad action, such as tax evasion (Prinz 2011a, 220). These cases strongly suggest that empathy can't be causally necessary for moral judgment. Indeed, they support a simpler explanation, which Prinz puts as follows:

My moral response is linked to action-types. If I classify your behavior as an instance of “stealing,” then that is enough to instill moral ire. Disapprobation can follow directly from certain types of action without any need to contemplate the suffering of victims. (2011a, 220)

This simple alternative hypothesis appears to be superior to any Empathy Causation account. But it requires an answer to the question of how we come to disapprove of actions of certain types, and there empathy might yet have a role to play.

2. Empathy Constitution

Could empathy be *constitutively* involved in each and every moral judgment? If it were, the following would be true. Here is the thesis:

Empathy Constitution

Moral judgments are constituted by affective empathy either with the patient or the agent of the action.

In its patient-focused form, Empathy Constitution is vulnerable to many of the challenges faced by Empathy Causation, or their analogues. For example, if we consider ourselves to be wronged, but can't empathize with ourselves, clearly our judgment isn't constituted by our empathic emotion. How about the agent-focused variant? It has recently been defended by Michael Slote. Slote begins by observing that "empathic concern for others is itself a psychological state that may be the subject or object of empathy" (Slote 2010, 34). Suppose that someone goes out of their way to help the homeless, thus manifesting a high degree of empathic concern. In Slote's terminology, she displays warmth and tenderness towards the homeless. If I come to share her feelings, I feel warmth and tenderness towards her, and such "empathy with empathy" constitutes moral approval (ibid., 35). On the other hand, if someone acts towards others with cold indifference, taking on that person's feeling means that I have a cold feeling towards the agent, and that constitutes my disapproval, according to Slote. Such emotional approval and disapproval then "enters into making moral judgments" (53), which helps explain why moral judgments coincide with motivation.

Many philosophers have criticized Slote's proposal (see e.g. Stueber 2011). One problem that I'll leave aside here is a more general metaethical issue of whether moral judgments could consist even in part in feelings of any kind, given their semantic and inferential properties. Apart from this issue, one obvious concern is that if I take on your warm feeling towards the homeless, I seem to end up with a warm feeling towards the homeless, not towards you. Slote's response to this is to claim that the intentionality of the empathic feeling is determined by its causal origin, which is the agent's feeling (2010, 39). But this view of the intentionality of empathic emotions has odd implications. Suppose you're angry with Eilis, and I empathize with you. Since the cause of my empathic anger is your feeling, Slote's view implies that I'm now angry with you! This is unacceptable.

Second, feelings of warmth or chill towards someone seem to have very different characteristics than moral approval or disapproval (Prinz 2011a). For example, while feeling cold toward someone is no doubt a negative feeling, it only accidentally motivates us to impose sanctions on the agent. It contrasts with blaming attitudes like resentment or indignation, which are hardly phenomenally ‘cold’ in any sense. Finally, agential empathy seems to be neither sufficient nor necessary for approval, nor its absence for disapproval. For example, we can disapprove of actions that are not done out of unempathic motives, such as an animal rights activist throwing a cake in the face of a politician out of empathy for the suffering of farm animals (cf. D’Arms 2011). In short, neither the patient- nor the agent-focused variant of Empathy Constitution is particularly plausible.

3. Empathy’s Role in Explaining Moral Norms

Given the challenges to Empathy Causation and Empathy Constitution, one might think that empathy can’t play a role in explaining moral judgment. But what has been said so far doesn’t yet rule out a more indirect, yet still necessary role for empathy in the *development* of moral judgment. Consider Prinz’s hypothesis that “moral response is linked to action-types” (2011a, 220). As I noted, it requires an explanation of how such a link is formed. That is, why do we embrace certain *specific* moral norms or rules, and why do we regard certain norms as specifically *moral*, as opposed to conventional? Here’s one hypothesis:

Empathy Explanation Hypothesis

Empathic feelings are a necessary part of the best explanation of a) why subjects endorse pro-social norms b) in a distinctively moral way.

By ‘pro-social norms’ I mean norms that prohibit harming other people in certain contexts, require respecting people and their property, demand fair treatment, and so on. Not all norms

that people have historically endorsed in a distinctively moral way are pro-social in this sense. Obviously, practices like slavery or marital rape have been considered morally permissible, and some people think chauvinistic patriotism is morally required. The best explanation for the existence of such norms will likely appeal to the self-interest of privileged populations rather than empathy. But empathy might play a role in explaining why people hold *some* central moral norms, even when it is against their self-interest or what they've been taught.

3.1 Empathy and Rules

The first part of the Empathy Explanation hypothesis is that empathy is necessary for explaining why people form judgments regarding certain pro-social act-types. There's several ways in which this explanation might work. Hume and Smith appealed to what might be called our *induction disposition*: once we perceive a pattern across cases, we project it to future instances as a generalized expectation. As Smith puts it:

The general maxims of morality are formed, like all other general maxims, from experience and induction. We observe in a great variety of particular cases what pleases or displeases our moral faculties, what these approve or disapprove of, and, by induction from this experience, we establish those general rules. (TMS 377)

It is, as I said, relatively uncontroversial that we do sometimes form judgments as a result of empathizing with someone. Suppose that during our formative years, we encounter several people who have been cheated by someone else, empathize with their anger or hurt, and consequently disapprove of the people who cheated them. (Insofar as we empathize impartially, we'll disapprove of cheating anyone, not just cheating people like us.) If we

have the induction disposition, we'll come to disapprove all instances of cheating by default, without having to empathize in each and every case of cheating.

A different but potentially complementary kind of empathy-based explanation says that over time, we (collectively) come to embrace pro-social moral norms out of a set of candidates, because those norms resonate with our empathic tendencies. Individuals might then pick up these norms through socialization without themselves empathizing. This would be an empathy-based variant of the view that Shaun Nichols (2004) has developed, according to which norms that match our affective reactions enjoy greater “cultural fitness” than those that don't, and thus get transmitted from generation to generation. Nichols's own account appeals to a “Concern Mechanism”, which is triggered by attributions of negative hedonic or affective states to others, and generates concern for them, reactive distress, and contagious distress (2004, chapter 2). It is these emotional responses that explain why norms against causing harm to others are widely adopted, according to Nichols. However, our actual pro-social moral norms are much more complex than blanket prohibitions against harming people – it matters to us what the agent's motives are, whether the harm is a means to an end or a side effect, whether the harm is the result of an action or of an omission, whether the harm is intended, negligent, or merely accidental, and so on (for some empirical data, including data about cultural variation, see Young and Tsoi 2013 and Barrett et al. 2016). Arguably, a simple Concern Mechanism cannot account for the greater cultural fitness of such norms that are sensitive to the agent's quality of will. Instead, explaining them may require appealing to perspective-taking along the lines of the regulated empathy hypothesis (see Kauppinen forthcoming for some details).

How successful is this part of Empathy Explanation? Properly answering this question would require a detailed comparison with alternative explanations of why we embrace pro-social moral norms. While philosophers sometimes appeal to the ability to

intuit moral principles or pure practical reason, psychologists tend to prefer evolutionary explanations that appeal to innate affective dispositions to disapprove of behavior that reduces fitness at the group level (Haidt 2012) or an innate moral ‘grammar’ that generates moral principles (Mikhail 2011). Jesse Prinz (2011a, 2011b) offers parental conditioning and imitation as an alternative. However, such a story is evidently incomplete, since it leaves unexplained why *parents* endorse and transmit some norms and not others. In any case, if any of these hypotheses is true, empathy is not necessary for explaining our adherence to our moral norms. What Empathy Explanation has going for it is parsimony, since it doesn’t require assuming any kind of innate moral capacity (see Nichols 2005). It also predicts that empathic reactions to novel situations will result in specification or rejection of pre-existing principles (Masto 2015) – for example, we might come to rethink our convictions regarding slavery or treatment of refugees as a result of vivid descriptions that arouse empathic feelings.

2.2 Empathy and the Moral/Conventional Distinction

The second part of Empathy Explanation is that empathy is necessary for explaining why we regard certain pro-social norms as distinctively *moral*. Clearly, not all norms belong in this category. We can, for example, think that something is against the law without thinking that it is morally wrong. What is distinctive of moral norms, then? Since the work of Elliot Turiel (1983), it has been common for psychologists to focus on different ways of responding to transgressions of norms. According to this tradition, some transgressions are regarded as wrong independently of whether they are permitted by social, political, or even religious authorities (authority-independence), as wrong everywhere (universality), as more seriously wrong (seriousness), and more severely punishable. These norm-violations include what are often considered paradigmatic moral wrongs, such as injustice or harming others. Other

transgressions, such as dressing in a particular way, are regarded as wrong only when prohibited by some local authority, such as a teacher, legislator, or custom, wrong only locally, as less seriously wrong, and less severely punishable. These features are taken by many psychologists to mark what is called the moral/conventional distinction. Many studies have found that children distinguish between these two kinds of transgression from an early age, roughly 2 to 3 years old (Smetana 1981). Although this distinction has recently become controversial (see Kelly et al. 2007 and Shoemaker 2011), the critiques arguably misconstrue what authority-independence in the relevant sense entails, so I will assume in the following that it is nevertheless along the right lines. (Space constraints prevent a more detailed examination here.)

Supposing this is the right way to draw the moral/conventional distinction, what could be the role of empathy in explaining it? Start with the contrast between a paradigmatic moral transgression that involves one person deliberately harming another in order to further their own perceived interests, and a paradigmatic conventional violation, such as wearing different colored socks. In the first case, if we either affectively empathize with the person harmed, or imaginatively put ourselves in her position, we will predictably have a negative reactive attitude such as resentment towards the agent. Typically, at least as far as we ourselves see it, this attitude doesn't depend on our personal relationship to either the agent or the patient, or on taking ourselves to be more important than other people, so we take it that any normal, decent person would feel the same way. Consequently, we emotionally construe the action as being wrong, whether or not we've been told by someone that it is impermissible (see Kauppinen 2013). Further, we're in a position to appreciate *why* the action is wrong – say, that it manifests insufficient regard for the victim's will or interests. And when we have a negative reactive attitude towards the agent, we're already blaming them, and if we take it that any informed and impartial spectator would feel the same way,

we already construe blame as fitting. In the second case, we won't have these emotional responses towards different colored socks. So even if both harming and dressing in a certain way go against rules that we've been taught, it's no surprise that we regard the first violation to have a different status, and consider it to be wrong regardless of whether someone in a position of authority permits it.

These considerations suggest that empathizing of a certain kind is *sufficient* to distinguish moral from conventional norms. It's a much more demanding task to make that empathy is *necessary* for recognizing the distinction. One challenge is that other emotional responses, such as disgust, seem sufficient to get some people to regard certain transgressions, such as masturbating with a dead animal or spitting into a glass from which one is going to drink, as authority-independently wrong (Nichols 2004). One line of response would be to emphasize that authority-independence isn't the sole mark of a moral norm. There's also the fittingness of blame and guilt. A defender of Empathy Explanation might insist that when we think that a disgusting behavior is morally wrong, there is an element of cognitive empathy involved in our disapproval – we imagine any normal person would blame the agent for such behavior. It would support this hypothesis if people who lack empathy couldn't genuinely distinguish between moral and conventional violations. I'll turn to this issue next.

2.3 Empathy Deficits and Moral Judgment

Empathy Explanation predicts that people with empathy deficits should manifest deficient moral judgment. (I will restrict my attention to judgment, not moral agency in general). In this context, two populations, psychopaths and autists, have received particular attention. In the following, it is worth bearing in mind that both conditions are spectrum disorders – any deficits associated with them can be expected to be a matter of degree.

Let's start with psychopaths. One characteristic of the disorder is that psychopaths care little or nothing about how other people feel. Nevertheless, psychopaths appear to be good at attributing feelings to others, possibly by way of perspective-taking. Thus, they seem to be deficient in affective empathy in particular, and seemingly a good test case for its necessity for moral judgment. Are psychopaths capable of moral judgment, then? This is controversial. They do well in some tests of moral reasoning, and generally classify as wrong the same actions as normal people do. But since they appear to be unmoved by the wrongness of certain actions, some philosophers deny that they genuinely consider them to be morally wrong – rather, they're only parroting what they've been taught. Whether this is the case hangs on whether moral judgment internalism is true or not (for the current state of the debate, see Björnsson et al. (eds.) 2015).

A metaethically more neutral test is whether psychopaths can distinguish between moral and conventional violations. Some evidence suggests that they can't – in particular, one study found that convicted criminal psychopaths say all transgressions are authority-independently wrong (Blair 1995). R. J. R. Blair's (1995) interpretation is that psychopaths can't tell the difference between moral and conventional norms since they lack an affective response to the suffering of others (so, unsurprisingly, they rarely appeal to harm as a justification for why a violation is wrong), but try to create a good impression in the eyes of authorities by erring on the side of caution and classifying all violations as authority-independent. However, more recent studies suggest that psychopaths may, after all, be able to make the distinction. For example, when Aharoni et al. (2012) gave a forced-choice test, in which subjects had to classify 8 out of 16 transgressions as authority-independent, they "found no evidence that high-psychopathy offenders—as measured by total psychopathy score—were any poorer at distinguishing moral from conventional transgressions than were low-psychopathy offenders." However, total psychopathy score includes a number of

components, such as interpersonal and lifestyle facts, in addition to affect, and they found that the affective facet taken alone *did* influence performance. (For an alternative interpretation of the data, see Levy 2014.)

In brief, then, evidence from psychopaths must be regarded as inconclusive at the moment. Matters are further complicated by the fact that psychopaths suffer from other deficits as well. Prinz (2011a) argues that their generally low affect levels suffice to explain any deficiencies in moral judgment, while Jeanette Kennett (2002) argues that it is problems with reasoning and impulse control that are the cause – in particular, she claims psychopaths are unable to appreciate how considerations independent of one’s present desires provide reasons that extend over time (2002, 355). In the latter vein, Heidi Maibom draws on various empirical studies to support the hypothesis that psychopaths have deficient practical rationality in the sense of willing means to their ends and ensuring that their aims are consistent, among other things, due to “impairments in attention width and span, impulsivity, deficient self-understanding, and difficulties adjusting their responses” (Maibom 2005, 253–254).

However, Aaltola (2014) observes that so-called secondary psychopaths are ‘hot-headed’ and aggressive, though not empathic, so Prinz’s hypothesis doesn’t seem to work for them. In contrast, primary psychopaths are extremely controlled and intelligent, while being emotionally detached and fearless. These ‘snakes in suits’, as they’re sometimes called, don’t seem to have the problems with reason that Kennett’s and Maibom’s hypotheses require (although such individuals have not yet been studied as carefully as incarcerated psychopaths). So it seems empirically plausible that it is the common problem with both kinds of psychopath, lack of empathy, which *best explains* the deficits in their moral judgment. But any definite conclusions would be premature.

Autistic people are another population of interest, since, roughly speaking, it is characteristic of them that they can't adopt the perspective of others, and thus rate low on cognitive empathy. At the same time, autists are often conscientious, and both autistic children (Leslie et al. 2006) and adults (Zalla et al. 2011) distinguish between moral and conventional norms. Kennett concludes from these facts that "the case of autism shows that both selves and moral agents can be created in the absence of empathy" (2002, 357). But this conclusion, too, may be premature. Many researchers believe that autists are capable of *affective* empathy, even if they are bad at mindreading (Dziobek et al. 2008), so a version of Empathy Explanation might still be true (Blair 2005). However, there is at least some reason to think that autism involves deficient affective empathy as well. Hobson and Hobson (2014) draw on various studies to argue that since the cognitive deficits of autists make it hard for them to experience others as persons with minds in the first place, the range, depth, and likelihood of their emotional response to the feelings or situations of others is severely limited.

Adding to the complexity of the issue, some recent research has called into question the assumption that autists are capable of making the moral/conventional distinction. In particular, Zalla et al. (2011) gave autistic subjects not only questions regarding standard moral and conventional violations, but also disgust violations drawn from Nichols (2004). What they found was that while normal subjects regarded moral and disgust violations as authority-independent, they nevertheless distinguished between them. But autists didn't. They also rarely appealed to the welfare of the victim to justify their judgment in moral cases. Other studies have shown that autists' violation judgments aren't as sensitive to agent's intentions as normal subjects' judgments – roughly, they judge unintended harms to be as bad as deliberate harms (Moran et al. 2011, Buon et al. 2013). These results are unsurprising – given the trouble autists have in putting themselves in other people's shoes,

they can be expected to have difficulty appreciating the moral significance of the agent's and patient's perspective on an action. On the basis of such considerations, Tiziana Zalla and co-authors conclude:

We argue that while the affective component of the empathy is sufficient to distinguish affect-backed from affect-neutral norms, an intact cognitive empathy, which is specifically involved in moral appraisal, is required to distinguish moral from disgust violations. (Zalla et al. 2011, 123)

This is good news for Empathy Explanation, since it suggests that cognitive empathy may after all be necessary for being able to properly make the moral/conventional distinction, possibly by way of making possible more cognitively demanding forms of affective empathy.

To sum up, although the empirical evidence regarding psychopathy and autism is controversial, it seems plausible that while members of both empathy deficient populations may be capable of distinguishing between moral and conventional violations in at least some cases, they have a poor grasp of the grounds for authority-independent rules for blaming people (cf. McGeer 2008 and Shoemaker 2015). This suggests that both cognitive empathy and affective response to putting oneself in the shoes of others may be necessary for *moral insight* and perhaps *moral intuition*. This more modest hypothesis predicts that empathy deficient people, regardless of their reasoning capacity, will be poor at making moral judgments when moral insight is needed – in particular, when the rules that one has learned from others don't yield an answer, or yield answers that conflict with one another. While it has not, to my knowledge, been empirically tested yet, this hypothesis does fit well with observed behavior.

4. Is Empathy Good or Bad for Moral Judgment?

Regardless of how empirical controversies regarding the causal or explanatory role of empathy turn out, we can ask whether empathizing with others before making moral judgments makes it more likely that those judgments are correct. In one kind of situation, this is the case rather trivially. The correctness of some moral judgments hangs on facts about the feelings and other psychological states of other people. For example, whether Emily is to blame for hurting Joe's feelings depends in part on what Emily's intentions were and whether Joe's feelings were indeed hurt. Assuming that cognitive empathy is one way of learning about other people's intentions and feelings, it will in such cases be conducive to making correct judgments. Further, we occasionally ask each other to "walk a mile in our shoes" before blaming us. In this kind of case, too, cognitively empathizing with someone is likely to change our moral responses for the better – typically, it leads us to better appreciate the presence of excusing factors. (The data from autists supports this hypothesis.) So, in brief, cognitive empathy with the agent or those affected by an action can be expected to improve moral judgment whenever the moral status of the action depends on empirical facts about mental states that are accessible via empathy.

But what if our moral judgments result from affective empathy with the actual feelings of other people? Is that a good thing? It is a commonsense notion that empathy serves as a check to bias and self-interest, so we should try to be more empathic if we can before we judge. But recently, some philosophers and psychologists have argued that empathy is inherently morally problematic. Paul Bloom (2013) maintains that it is "parochial, narrow-minded, and innumerate", so we shouldn't rely on it. He notes that psychological research has revealed what is called the identifiable victim effect: people seem to care more about the plight of one individual they can relate to than about the suffering of many people that shows up in statistics (Jenni and Loewenstein 1997). Indeed, we can't possibly empathize with everyone, and people who are strangers to us might be particularly difficult

for us to empathize with, so we're better off relying instead on "more abstract principles of justice and fairness, along with a more diffuse compassion" (Bloom 2015). Sometimes this means going against empathy's verdict, as when punishment is warranted, or fair allocation of resources means that someone has to suffer. Jesse Prinz (2011a, 2011b) makes similar criticisms, adding that empathy is easily manipulated (for discussion and a response to Prinz, see the Shoemaker chapter). We might also note that since it can be difficult for the privileged and powerful to put themselves in the shoes of the underprivileged – the poor, the disabled, or ethnic minorities, for example – relying on empathy in a political context may lead to reinforcing the unjust status quo.

With the notable exception of Michael Slote (2010), who thinks moral demands are as partial as empathy is, most partisans of empathy find these observations troubling. Alas, these concerns are not new. As already discussed, none other than David Hume himself was keenly aware of the biases and limitations of our empathy. We've seen that he believed we can, to some extent, correct for the inbuilt biases. Nevertheless, he acknowledged that empathy ("humanity") can easily come into conflict with justice:

When I relieve persons in distress, my natural humanity is my motive; and so far as my succour extends, so far have I promoted the happiness of my fellow-creatures.

But if we examine all the questions, that come before any tribunal of justice, we shall find, that, considering each case apart, it wou'd as often be an instance of humanity to decide contrary to the laws of justice as conformable to them. Judges take from a poor man to give to a rich; they bestow on the dissolute the labour of the industrious; and put into the hands of the vicious the means of harming both themselves and others. (T 155)

So far, Hume is in agreement with critics like Bloom. But he digs deeper. Hume believes that justice is an "artificial virtue", a system of rules that has arisen, because it serves, on the

whole, each individual's enlightened self-interest. But why should we think that serving not just our own interest but the interests of all (or at least most) others is good or just? Hume's answer is that it's because we empathize with those others, and thus disinterestedly approve of justice and other artificial virtues. Given his view of the nature of empathy, Hume thought this would result in endorsing a kind of rule-utilitarianism. Adam Smith, in contrast, held that since we might empathize with the resentment of one person sacrificed for the benefit of many even if we're impartial, empathy-based rules are non-consequentialist.

Either way, such indirect empathy is arguably a good thing, when it comes to settling on principles of justice – someone who lacked such empathy might end up endorsing rules that fail to serve the general good, or show insufficient regard for the dignity of each individual. Consider someone like Adolf Eichmann, who possibly quite sincerely thought that morality required him to do his duty as defined by his superiors (Arendt 1963). It is safe to assume that had he reflected on the rules he was told to obey by placing himself of in the position any one of those negatively affected them, he wouldn't have been equally enthusiastic about obedience to Nazi authorities. After all, empathy, in particular when it is regulated by reference to an ideal of impartiality, tends to result in embracing pro-social norms of the kind that most of us regard as correct (cf. Hoffman 2011).

So, philosophers who think empathy is a good thing for morality have always been aware that our natural tendency to take on other people's feelings is an unreliable guide in moral judgment. Instead, they have argued that empathy must be tempered by regulating our emotional responses to particular cases so that we won't miss the big picture. Nevertheless, they maintain that the alleged conflict between empathy and justice is illusory, since our regard for justice demands itself stems from empathy. They would thus agree that we shouldn't base decisions about immigration policy on how we feel when contemplating a dead boy, for example. Untutored empathy can blind us to the non-actual and the wider

context. But it can spur us to reflect on what the alternatives to the actual situation are, and to consider their impact in the light of principles that may get their grip on us in virtue of resonating with impartially empathic responses.

4. Conclusion

It is likely an exaggeration to claim that empathy is the “cement of the moral universe”, as Michael Slote (2010) does. It is not plausible that empathy is either causally or constitutively involved in each and every moral judgment, and it may not be necessary for having the capacity to distinguish moral from other norms. But people who lack the ability to put themselves in the place of others and feel for them do appear to have trouble with moral insight and appreciating the grounds of pro-social moral principles, even if their rational powers are largely intact. This suggests that empathy may have an irreplaceable role in the development of good moral judgment after all, although it wouldn’t be wise to rely on it in each individual case.

Bibliography

- Aaltola, Elisa 2014. Affective Empathy as Core Moral Agency: Psychopathy, Autism and Reason Revisited. *Philosophical Explorations* 17 (1), 76–92.
- Aharoni, Eyal, Sinnott-Armstrong, Walter, and Kiehl, Kent A. 2012. Can Psychopathic Offenders Discern Moral Wrongs? A New Look at the Moral/Conventional Distinction. *Journal of Abnormal Psychology* 121 (2), 484–497.
- Arendt, Hannah 1963. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Viking Press.
- Blair, R. J. R 1995. A Cognitive Developmental Approach to Morality: Investigating the Psychopath. *Cognition* 57, 1–29.

- Bloom, Paul 2013. The Baby in the Well: The Case Against Empathy. *New Yorker*, May 20, 2013.
- Bloom, Paul 2015. Imagining the Lives of Others. *New York Times*, June 6, SR8.
- Björnsson, Gunnar, Björklund, Fredrik, Strandberg, Caj, Eriksson, John, and Ragnar Francén Olinder 2015 (eds.). *Motivational Internalism*. Oxford: Oxford University Press.
- Buon, Marine, Dupoux, Emmanuel, Jacob, Pierre, Chaste, Pauline, Leboyer, Marion, and Zalla, Tiziana 2013. The Role of Causal and Intentional Judgments in Moral Reasoning in Individuals with High Functioning Autism. *Journal of Autism and Developmental Disorders* 43, 458–470.
- Coplan, Amy and Goldie, Peter (eds.) 2011. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- D’Arms, Justin 2011. Empathy, Approval, and Disapproval in *Moral Sentimentalism*. *Southern Journal of Philosophy* 49, Supplementary Volume, 134–141.
- Dziobek, Isabel, Rogers, Kimberley, Fleck, Stefan, Bahnemann, Markus, Heekeren, Hauke R., Wolf, Oliver T., and Antonio Convit 2008. Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders* 38, 464–473.
- Haidt, Jonathan 2012. *The Righteous Mind*. New York: Pantheon Books.
- Hobson, R. Peter, and Hobson, Jessica A. 2014. On Autism: A Perspective from Developmental Psychopathology. In Maibom (ed.) 2014.
- Hoffman, Martin 2011. Empathy, Justice, and the Law. In Coplan and Goldie (eds.) 2011, 230–254.
- Hume, David (1739–40/1978) *A Treatise of Human Nature*. Ed. L. A. Selby-Bigge, 2nd rev. edn., P. H. Nidditch. Oxford: Clarendon Press. (T)

Jenni, Karen and Loewenstein, George 1997. Explaining the “Identifiable Victim Effect”.

Journal of Risk and Uncertainty 14, 235–257.

Kauppinen, Antti 2010. What Makes a Sentiment Moral? *Oxford Studies in Metaethics* 5,

225–256.

Kauppinen, Antti 2013. A Humean Theory of Moral Intuition. *Canadian Journal of*

Philosophy 43 (3), 360–381.

Kauppinen, Antti 2014. Empathy, Emotion Regulation, and Moral Judgment. In Maibom

(ed.) 2014, 97–121.

Kauppinen, Antti forthcoming. Sentimentalism, Blameworthiness, and Wrongdoing. In

Remy Debes and Karsten Stueber (eds.), *Ethical Sentimentalism*. Cambridge:

Cambridge University Press.

Kelly, Daniel, Stich, Stephen, Haley, Kevin J., Eng, Serena J., and Fessler, Daniel M.T. 2007.

Harm, Affect, and the Moral/Conventional Distinction. *Mind and Language* 22 (2),
117–131.

Kennett, Jeanette 2002. Autism, Empathy, and Moral Agency. *Philosophical Quarterly* 52

(208), 340–357.

Leslie, Alan, Mallon, Ron, and DiCorcia, Jennifer 2006. Transgressors, Victims, and Cry

Babies: Is Basic Moral Judgment Spared in Autism? *Social Neuroscience* 1 (3–4),
270–283.

Levy, Neil 2014. Psychopaths and Blame: The Argument from Content. *Philosophical*

Psychology 27 (3), 351–367.

Maibom, Heidi 2005. Moral Unreason: The Case of Psychopathy. *Mind and Language* 20

(2), 237–257.

Maibom, Heidi (ed.) 2014. *Empathy and Morality*. New York: Oxford University Press.

- Masto, Meghan 2015. Empathy and Its Role in Morality. *Southern Journal of Philosophy* 53 (1), 74–96.
- McGeer, Victoria 2008. Varieties of Moral Agency. Lessons from Autism (and Psychopathy). In Walter Sinnott-Armstrong (ed.), *Moral Psychology Vol. 3: The Neuroscience of Morality*. Edited by Walter Sinnott-Armstrong. Cambridge, MA: MIT Press, 227–257.
- Mikhail, John 2011. *Elements of Moral Cognition*. Cambridge: Cambridge University Press.
- Moran, Joseph, Young, Liane, Saxe, Rebecca, Lee, Su Mei, O’Young, Daniel, Mavros, Penelope, and Gabrieli, John 2011. Impaired Theory of Mind for Moral Judgment in High-Functioning Autism. *PNAS* 108 (7), 2688–2692.
- Nichols, Shaun 2004. *Sentimental Rules*. New York: Oxford University Press.
- Nichols, Shaun 2005. Innateness and Moral Psychology. In Peter Carruthers, Stephen Laurence and Stephen Stich (eds.), *The Innate Mind: Structure and Contents*. New York: Oxford University Press, 353–369.
- Prinz, Jesse 2011a. Against Empathy. *Southern Journal of Philosophy* 49, Supplementary Volume, 214–233.
- Prinz, Jesse 2011b. Is Empathy Necessary for Morality? In Coplan and Goldie (eds.) 2011, 211–229.
- Shoemaker, David 2011. Psychopathy, Responsibility, and the Moral/Conventional Distinction. *Southern Journal of Philosophy* 49, Supplementary Volume, 99–124.
- Shoemaker, David 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.
- Slote, Michael 2010. *Moral Sentimentalism*. New York: Oxford University Press.
- Smetana, Judith 1981. Preschool Children’s Conceptions of Moral and Social Rules. *Child Development* 52 (4), 1333–1336.

Smith, Adam 1759–1790/2002. *The Theory of Moral Sentiments*. Ed. Knut Haakonssen.

Cambridge: Cambridge University Press. (TMS)

Stueber, Karsten 2011. Moral Approval and the Dimensions of Empathy: Comments on

Michael Slote's Moral Sentimentalism. *Analytic Philosophy* 52 (4), 328–336.

Turiel, Elliot 1983. *The Development of Social Knowledge: Morality and Convention*.

Cambridge: Cambridge University Press.

Young, Liane and Tsoi, Lily. When Mental States Matter, When They Don't, and What That

Means for Morality. *Social and Personality Psychology Compass* 7 (8), 585–604.

Zalla, Tiziana, Barlassina, Luca, Buon, Marine, and Leboyer, Marion (2011). Moral

Judgment in Adults with Autism Spectrum Disorders. *Cognition* 121(1), 115–126.