PHIL 640: Empathy

Brian Weatherson

November 22, 2021

What is Empathy

Most weeks when we're doing one of these things focussed on a particular virtue, we start with questions about what it is, and why it might or might not be a virtue. We're not going to quite do this here, because it seems common ground that there are a lot of different things that we might mean by empathy, and it doesn't really matter which of them the ordinary term refers to. Still, it's worth clearing a few things out first.

First question, which kind of state is core to empathy?

1. Evaluative
2. Cognitive
3. Affective

By 'evaluative', I mean valuing the interests and well-being of others. That doesn't seem like nearly enough for empathy. It isn't a particular virtue, but the basis of morality or something close to it. Let's set that aside.

Most of the writers we're interested in don't think merely cognitive empathy is enough to count, but this is a bit more interesting. By cognitive empathy, I mean that when someone is in a particular (morally significant) way, the cognitively empathic person knows that they are that way. So if A is sad, B knows that A is sad (or something like this). And hopefully, if B is evaluatively empathic, B has a reason to do something about this.

The central notion of empathy we're interested in, and again this is more stipulation than analysis, is the affective notion. When A feels a certain way, the affectively empathic person B will also feel that way.

Cross-cutting this three-way distinction, we can also ask whether empathy requires something like exact match, or just something (morally) close enough. This isn't a binary distinction, 'close enough' is obviously gradational. But it's helpful to simplify and think about the more and less extreme versions. We'll call these the exact and inexact versions of the theory.

So on the exact version, if A is bored, then exact, affective, empathy requires B to feel bored (in the same way). But inexact, affective empathy requires B to feel something close to that. Maybe feeling upset or angry that A is bored is enough. And inexact cognitive empathy might require merely that B knows that A feels badly in some way or other.

If we were trying to do conceptual analysis of the ordinary concept EMPATHY, we might wonder next

about whether some conjunctive or disjunctive combination of these is the right account. Maybe ordinary empathy requires exact evaluative and cognitive empathy and inexact affective empathy. But let's not go there, because we're really interested in exploring the conceptual space, not in matching up that space to ordinary words. And noting these six basic ideas (three kinds of empathy crossed with exact vs inexact) seems like enough.

One thing I will note is that some of Bloom's complaints, perhaps all of the important ones, are targeted at exact affective empathy. He doesn't think it's part of virtue that one is bored when people around you are bored, or depressed when people around you are depressed. And I'm inclined to agree - that doesn't seem like it's much good to anyone. (Though we'll come back to reasons to worry about this.) But there are lots of other notions of empathy.

Virtues and Ability

There is something that it's perhaps worth pausing over here, though it's a bit delicate and I don't know quite what to say about it. The following three claims are in some amount of tension.

1. The six notions of empathy we started with (or at least four or five of those six) are virtues.
2. Many of these six notions represent attributes that it will be impossible, on broadly medical grounds, for people with certain disabilities to have.
3. Having a disability never makes one less virtuous.

The general structure of the tension here is something that worries me a lot in other contexts. I'm sympathetic to the Murdoch view that a core part of morality involves seeing things in the right way. There are a couple of ability-related reasons that you might worry about this view. One is that the language is bad - you don't want to use language that draws an analogy between good people and sighted people as opposed to blind people. And that's fair enough, though if that's all that's going on we should complain about the metaphor and work on fixing the language. But there's a second deeper complaint. Maybe the view doesn't just draw an analogy between physical sight and virtue, but genuinely makes physical sight an aspect of virtue. And that really seems like we're slipping onto uncomfortable territory.

And something like the same thing happens with cognitive and affective empathy. The ability to know what others are feeling, and to feel the same way, is an ability. That's to say, sometimes people don't have that ability because they are disabled. That's fine, not everyone can do the same things, and we don't normally make that morally significant. Saying that someone is a less virtuous person because they don't have this ability (to the same extent) is worrying. But I don't really have a good story here. It's just something to keep in mind going forward.

Empathy and Equality

So here a couple of reasons for being sceptical of empathy in ethics.

1. People who use empathy in forming judgments about what to prioritise do not make the judgments a utilitarian would approve of.
2. The right judgments are those that a utilitarian would approve of.
3. So it's bad to use empathy in forming judgments about what to prioritise.

So obviously most people won't like premise 2 in that. (Though not everyone!) When Bloom (and people like him) talk about empathy leading to "innumeracy", I sometimes worry about the implicit argument being something like this one. Could there be something better than could be meant?

There is, I think, a more compelling argument in the vicinity. I'm not going to make this valid, but you can see how to fill in the gaps I hope.

1. It's impossible, certainly in practice and maybe in theory, to empathise equally with everyone.
2. In practice (and maybe in theory) we empathise more with those who are already relatively privileged.
3. If we use empathy in forming priority judgments, we'll prioritise those we empathise more with.
4. It's bad to not treat people equally when forming priority judgments, and it's especially bad to violate equality by favoring those who are already privileged.
5. So, it's bad to use empathy in forming priority judgments.

The 'in practice' part of premise 1 should be obvious. But I think it's a little interesting to think about the 'in principle' part. Empathy is sharing the experiences of others. But I don't know how you share the experiences of many people at a time. It would violate the unity of consciousness. So there's a fairly deep sense in which being empathic is to privilege the one over the many.

Premise 2 should be reasonably clear as well, at least for people like us. We find it easiest to empathise with people like us. If nothing else, there are fewer epistemic barriers. And we are, at least in a global context, incredibly privileged. But I suspect there is more to it than that. Empathy involves focussing on the feelings of a particular person. And it's easier to single out a privileged person than an under-privileged one. Tolstoy's quip about families isn't really true about individuals - a lot of underprivileged individuals are underprivileged in depressingly similar ways. But privileged people get distinctive privileges. Jeff Bezos and Elon Musk have really different lives (even if both of them share an obsession for space travel).

Premise 3 seems relatively trivial, and hopefully premise 4 is plausible. You don't have to be a consequentialist to think that some basic principle of equality is central to ethics.

Still, while the premises don't look as utilitarian as the first argument, you might think the conclusion will still be fairly utilitarian. Or, at least, that it will lead to something like utilitarianism with side-constraints. Maybe that's ok, but it's worth thinking ahead to where this leads before you sign up for the premises.

Good Acts and Good People

One conclusion you might take from Bailey's paper is this:

- There are some goods that only bad (or at least suboptimal) people can bring about.

She presents this as a challenge within and for virtue ethics. But you might equally take it to be a problem for a certain kind of weak consequentialism. (This is, fwiw, how I take cases with this structure.)

Here's what I mean by 'weak consequentialism'. Consider the view that says it's more or less analytic that the right action is the one that produces the most good. This isn't to say the right action produces the most welfare, or surplus of pleasure over pain, or anything like that. It's just that rightness is the thing that's produced by good actions. And virtuous people, the view continues, just are those who do right actions.

The weak consequentialist, as I'm understanding them, doesn't commit to any order of explanation here. They are committed to the identity between right actions and actions typically performed by virtuous people, but not to any order of explanation. It might be, as Kant is typically read as holding, that the rightness of actions is analytically prior. Or it might be that the virtues of agents are prior, as modern virtue theorists think. Or it might be that goodness of outcomes is analytically prior, as 'stronger' consequentialists think. All they insist on is the tie between these notions.

Sometimes this kind of view can seem like it is hard to argue against. Perhaps you can argue against it by looking at various kinds of dilemmas, and how they play out on different views. If you're interested, the literature on this goes under the name 'consequentializing'. The person I'm calling a weak consequentialist says all views can be 'consequentialized'. And while not everyone agrees, a lot of people think this works, at least as a verbal trick.

Anyway, Bailey's examples give us a way to directly argue against such a view. Perhaps empathising with a bad person is something that (a) produces goods, and (b) is something only a (somewhat) bad person can do.

It's really crucial here that the empathy be partially constitutive of the good, not merely part of the typical causal background of the good. It's not a surprise that bad people might find it easier to produce certain goods, or would be more likely to produce certain goods. Here's one way that can happen. Imagine that X is bad not in that they value bad things, but in that they fail to value most good things. In particular, X only values one particular kind of good thing. (It won't matter what it is.) Now goods that involve promoting that one particular thing are more likely to be realised if X is around than if a good person is around. That's because the good person might think that promoting this one good has intolerable opportunity costs. But the monomaniacal X, who has only one interest, will definitely promote it. And we said the interest is something that's actually good. So there is a good they are more likely to bring about. I don't think that's a particularly interesting result. It's not surprising that when the right thing to do involves some conflict, having a bad person around who doesn't feel the force of the conflict could be useful. What would be surprising is if the good person was genuinely incapable of producing the good result. That's what Bailey's examples suggest.

You might wonder why this is a problem for the virtue theorist. After all, the virtue theorist isn't a consequentialist. Couldn't they join the deontologist in just saying that yeah, sometimes life is tragic, and following the rules will lead to bad results? Well maybe some virtue theorists could, but not the most common kind. Because the most common virtue theorists say that what makes something a virtue is that it is connected in the right kind of way to human flourishing. It would require root-and-branch repair of the theory if we discovered that some virtues were in fact inimical to human flourishing. And, again, that's what the examples suggest.

Virtue and Possibility

What gets the puzzle going is the thought that good people simply won't see bad actions as possible. There is something to this. I mentioned one example in class last week, but maybe it's worth writing down.

A has fallen asleep on a park bench, with their phone in their lap. It's right there to steal. I mean, it would

be really easy to steal. Two people walk by, B and C. B thinks about the phone, thinks about how easy it would be to steal and sell it, and then thinks, no that would be wrong. C looks at A and worries that the phone might get stolen because there are bad people around, but concludes that moving the phone to make it less visible would itself be a violation of A's space, so just hopes for the best. Happily for all concerned, A soon wakes up, and goes back to doomscrolling.

So question, which of these people are good? Well, don't worry about A - they were just asleep. But how should we think about B and C. Bailey is interested in a tradition, tracing to McDowell, that says only C is truly virtuous. B is merely continent.

As an aside, I don't love the visuals of the truly bad person as incontinent. It goes against the Austinian view that one can break the rules with some amount of style.

> I am very partial to ice cream, and a bombe is served divided into segments corresponding one to one with the persons at High Table: I am tempted to help myself to two segments and do so, thus succumbing to temptation and even conceivably (but why necessarily?) going against my principles. But do I lose control of myself? Do I raven, do I snatch the morsels from the dish and wolf them down, impervious to the consternation of my colleagues? Not a bit of it. We often succumb to temptation with calm and even with finesse. ("A Plea for Excuses", 198)

Anyway, back to our main plot-line. What should we say about B and C? McDowell says that C is the virtuous one. I suspect Murdoch should be credited here as well - let's come back to that. (I don't mean to criticise McDowell at all here; I think he is very up front about how he is following Murdoch.) Bailey worries about the following objection. No one is really like C. At least not always. Everyone is always a little tempted by the benefits that theft has over honest toil. So the result on this view will be that nobody is truly virtuous.

Armstrong used to say of certain theories he didn't like that they had the benefits of theft over honest toil. I used to always think that if that was the worst you could say about a theory, it might be time to start checking where the metaphysical security guards were.

At this point you might think that yes, that's definitely a thing that follows from the account of virtue. But is it anything more? That a theory of virtue implies there are no perfectly virtuous people seems no more a problem for it than it would be a problem if my theory of unicorns implies the world is unicorn free. Do we have wonderful independent evidence that the world is populated by moral saints? Are they hiding under the sofa? I don't know, but I'm not moved by this at all.

What I am moved by are two possibly related considerations. First, I think there's a sense in which B is praiseworthy and C is not. B overcame temptation. As Oscar Wilde almost quipped, that's the hardest thing to overcome. Good for B! Second, you might think, in a broadly Kantian spirit, that morality is all about intentional action. And C's not considering the possibility of stealing the phone might be a manifestation of a useful habit, but I don't think it's intentional. So if morality is all about intentional action, there cannot be a sense in which C is better than B.

So here's a big research question that I don't know the answer to. How much does this kind of reasoning drive Murdoch's anti-Kantianism? That's for another kind of seminar I'm afraid.

Silencing

What do McDowell, and Bailey, mean when they say that reasons to perform bad acts are *silenced*. They don't mean the kinds of thing people usually mean these days in philosophy by silencing. This isn't anything from the Hornsby-Langton school about how speech acts can't be recognised as such. What's silenced here is not a person but a reason.

As Bailey notes, there are a bunch of different looking things that McDowell might mean by this. But they are only different looking. They have different meanings on different ways to translate McDowell into more familiar Humean-speak. McDowell, however, is not a Humean. And sometimes that means if you try to specify a theory in Humean language and ask which bit he disagrees with, you won't get a satisfying answer. Because what you need to say is that the translation misses something.

This is a philosophical problem - how do we understand a disagreement between people when we struggle to articulate a thesis that one accepts and the other rejects? But it's also a practical problem for me, because I'm enough of a Humean that at some level what I'm doing when I'm reading non-Humean works is translating them into my language, and seeing what I disagree with. And that might be inappropriate here. But that said, I think I have some sense of what silencing means.

> Silencing is when an option is taken off the table because it is morally unacceptable. We don't compute the expected costs and benefits, goods and bads, rights and wrongs, of the action. We just don't consider it as one of the options.

I'm going to go with that interpretation, although you really should note that it's my best guess, and this is not really my field. One interesting consequence of this interpretation (or perhaps reason to think it isn't right) is that it isn't only reasons *for* doing immoral actions that are silenced. It's also reasons *against* doing them. Imagine I'm thinking of stealing a car. And there is a real downside to this, I might get caught. That's a reason for me not to steal the car. On my way of understanding the view, that reason is silenced for the virtuous agent. They just don't think about stealing the car.

For what it's worth, I sort of think this picture is built into the criminal law. It's true that the criminal law has punishments that are meant to disincentivise. But I think the picture the law has is that these are meant to be redundant for the good citizen. The good citizen doesn't weigh up the expected costs and benefits of stealing a car, and the role of retrospective punishment is not merely to alter that retrospective calculation. Rather, the good person hears the pronouncement *Thou shalt not steal cars* and doesn't steal cars. Of course, not all people are good people and the punishments have to be harsh enough to disincentivise the bad people. But that's a secondary role for the criminal law. End of digression on crime I guess.

But it's going to be important for what follows that we be careful about what exactly is silenced. Think about Bailey's example of the pears. On this picture, the virtuous person won't think of stealing the pears as even an option, so they won't think about reasons for or against stealing them. So far so good. But that doesn't mean they can't think about the pears. If you ask them "If, purely hypothetically, one could legitimately acquire these pears, they'd taste really good wouldn't they?", they would surely answer yes. Silencing doesn't mean you can't ask hypotheticals. It's just that if you ask them "If, purely hypothetically, I just stole these pears and ate them, I'd really enjoy them wouldn't I?", they'd faint or something. (I'm not sure what the intended causal mechanism is by which these reasons are silenced.)

The point here is that the virtuous person still knows that pears are tasty, and knows that is a reason to eat pears, and knows that this is a reason to possess pears. They just don't see it as a reason to *steal* pears. And this means that there are limits to the limits to their empathy. Although Bailey doesn't make a big point of this worry, I think she's sensitive to it, and that's why she constructs some of her examples the way that she does.

Understanding

What's the good thing that only bad people allegedly can do? It's to have a certain kind of understanding. Linking back to what we discussed at the start, there are two aspects I think to this kind of understanding.

1. One is to do with reasons. It's to appreciate the reasons that someone has, and to appreciate them as reasons. So it's not just to know that pears are tasty, but to appreciate that the tastiness of the pears is a reason for someone to act.
2. The other is to do with phenomenology. It's to know what it's like to be in the position that the person being understood is in. And that probably (maybe) requires having that experience, or at least something very much like it.

You might just about think that these two are closely connected. It's much easier to appreciate something as a reason if you know what it feels like to get it and to lose it. But we have to be a little careful here. The connection I just stated is causal, and what we're really interested in here are constitutive connections.

As noted earlier, it isn't a surprise that some bad people are more useful in some situations than good people. What would be surprising is if there is something a good person cannot do, just in virtue of their being good people. And here the question is whether they can provide the sort of empathic, humane understanding that friends need and perhaps deserve.

There are a couple of different cases that Bailey talks about, and which I had different reactions to: the pears and the schadenfreude. I wasn't really convinced by the pears example. True, the virtuous person (on McDowell's view of virtue) can't empathise with the desire to *steal* the pears. But they can empathise with the desire to *possess* the pears. They know that pears are tasty, they even know the way in which they are tasty, and they can appreciate that as a reason for wanting to possess them. It's a very strange kind of virtue that says that if in fact the only way to possess the pears is illegitimate, the virtuous person cannot even imagine legitimate possession, and the benefits that come with it. Imagination isn't bound by factual niceties like this, and imagination should be enough. So I don't really think that's a clear example. Maybe the virtuous person will find it harder to empathise, because they have to imagine the facts being a little different so the possession is legitimate, but that's not I think a big deal.

The schadenfreude example is much more interesting. Imagine A is virtuous, B is A's vicious friend, and C is B's enemy. B feels bad because something bad almost, but did not, happen to C. And they want A to empathise. And here I can see why it would be harder for the virtuous person. The first solution to the puzzle Bailey mentions, imaging that A could have an offline mind that can see what it's like to desire schadenfreude and an online part that doesn't, doesn't feel plausible. At least, I can't picture being so invested in the picture of virtue that makes this a problem and also feeling that this offline/online picture makes sense.

But maybe we should ask how big a deal it is if you need some bad people around for this. What's the cost of saying that only bad people can help you if you need someone to empathise over the failures of your bad schemes? Put that way, it sounds like something we should maybe expect. So let's leave it with that point - what's the downside of just having this outcome?