

Olivia Bailey

Empathy, Care, and Understanding in Adam Smith's *Theory of Moral Sentiments*

I. Introduction

When we observe and interact with other people, it is impossible for us to really think that they are not also experiencing the world. Inevitably, we understand people to be minded. And not only that: it also seems to us utterly obvious that we have some acquaintance with the content and character of others' experiences. I understand that the store clerk is bored, and that the children long for ice cream. I understand how the bitterly disappointed fourth-place finisher is feeling, and what the shame of the man who accidentally knocked over the Ming vase must be like. When we encounter another person whose inner world is a true mystery to us, it is a startling and disquieting experience. These individuals are the unsettling exception, rather than the norm. At the same time, it also seems obvious to us that we are interested in others' passions, thoughts, and attitudes, independent of their instrumental significance for us, and not only out of intellectual curiosity. That is, we care about others' inner lives. What others think and feel matters to us in a way that can directly motivate us to act on their behalf.¹

In *The Theory of Moral Sentiments*, one of Adam Smith's aims is to defend these commonsense convictions about care and understanding from solipsism and psychological egoism. Why does Smith see these twin positions as posing a threat that demands a response? At least part of the reason, I think, is that he endorses a particular picture of the difference between self-directed understanding and concern and other-directed understanding and concern. I will call this the 'picture of egocentric primacy.' This picture presents our self-care and self-understanding as uniquely primitive and self-evident. Our own minds are immediately transparent to us, but "we have no immediate experience of what other men feel" (TMS, I.i.1.2). Similarly, our self-care is immediate, but our care for others looks like the kind of thing that we can only come to through considerable mental maneuvering. Once one admits that our understanding and care for others needs to be explained, one had better follow up with an explanation. The skeptic can press the point: if the origin and nature of our understanding and care for others is

¹ Or, indeed, *against* their interests—care in the sense that I am concerned with is not always altruistically oriented.

mysterious, and if we cannot dispel that impression of mystery with a plausible explanation of how and why we come to care for and understand others, then our everyday impression that we do in fact understand and care for others may not be good enough reason to reject psychological egoism and solipsism.

One resource that has appealed to both early-modern and contemporary philosophers faced with this conundrum is the psychological phenomenon now generally known as empathy. What exactly empathy consists in has been the subject of lively debate since the time of Hume and Smith's exchanges about the nature of what they called "sympathy, or...fellow feeling" (TMS, I.i.1.3; see also TMS, VII.iii.3.17). For our purposes, though, the following loose but handily ecumenical definition of empathy should suffice: empathy consists in feeling what another person feels, or at least imagining feeling what another person feels, not because one is literally in the same situation as she is, but because one has in some other way come to imaginatively engage with something like the other person's experience. Empathy tends to interest philosophers operating with some version of the picture of egocentric primacy because it seems like it could provide the means of 'bootstrapping' up from our intimate understanding of and care for ourselves to our understanding of and care for others. In this paper, I will consider Smith's attempts to fend off skeptical worries about understanding and care by appealing to our capacity for empathy. My aim here is partly critical, and partly rehabilitative. Taken together, Smith's explicit pronouncements on empathy's role in fostering understanding and care generate a problem. Like his friend Hume, whose theory of empathy heavily influences Smith, Smith assigns a double duty to the series of mental operations that issue in empathy, treating it as the source of both our understanding of other people and our non-instrumental care for them. However, I will argue that if Smith's empathetic mechanism does generate an accurate understanding of the other, that understanding will simply not be the right kind of acquaintance to generate care. Only a seriously confused grasp of the attitudes and passions of the other could give birth to a heretofore absent non-instrumental care for others. As we will see, the fulfillment of either one of the empathetic mechanism's supposed functions requires conditions that will make it impossible for the other function to be fulfilled.

Smith's official account of empathy is seriously flawed. However, his theory of human sociability also contains within it the seeds of an important improvement upon the official account. Whether or not Smith actually recognizes it, one element of his account actually demands the conclusion that an important form of care precedes empathy and empathetically-derived understanding. Specifically, his explanation of why we engage in imaginative projection actually entails (or at least comes close to entailing) that concern for the other's thoughts and feelings is often (if not always) a necessary condition of our being motivated to empathize. This would mean that care is not something we get to through empathetic 'bootstrapping up' from self-care; it is prior to empathetic feeling. And if care is prior to empathetic feeling, then the problem that emerges from his official account does not get off the ground: that problem arises only because Smith officially treats both care and understanding as empathetically-derived. I will aim to make it clear that given Smith's conception of empathy, he should on pain of inconsistency be committed to a very different conception of care's relation to empathy and understanding than the one he more explicitly endorses.

II. Smithian empathy: an overview, and a problem

i. care and understanding

Before entering into the details of Smith's account, I want to identify more precisely the nature of the care and understanding that interests Smith. Let us first consider the range of attitudes we might classify under the heading of 'care'. What does it mean to care for or about someone? We tend to think of care as an attitude of benevolence, one associated with a motivation to aid the object of our care. However, we can also identify a broader sense of care, one that encompasses all of our interest in others and their inner lives, benevolent or not. Of course, our interest in others may be more or less motivated by our own self-concern. The kind of care that Smith regards as most in need of explanation is a concern for others' experiences, attitudes, emotions, and intentional actions that is not motivated by self-interest. I will refer to this particular kind of care as *basic care*. Significantly, this kind of care includes distinctly non-altruistic attitudes. For instance, I might care about my enemy's agony just because his unhappiness matters to me absolutely.

We can now turn to understanding. Understanding ranges along a scale of depth. At the shallow end, we have the understanding that consists in an acquaintance with another's behavioral patterns. This is mere predictive understanding. Then there are the varieties of understanding that are relevant only in those cases where we recognize the individual in question as having a mind. Classification of the various forms our understanding of other inner lives can take is inevitably contentious. One particularly difficult question is this: if we are to understand another's inner life, how important is it that we appreciate the other's experience? By experience, I mean the felt character of inner life, that aspect of our mental existence that we tend to refer to as 'what it is like.' One might think that we can distinguish between two ways of understanding another as minded, as follows: on the one hand, some understanding of an individual's feelings and beliefs can be reproduced in propositional form without remainder. And on the other hand, some understanding includes an ineffable sense of 'what it is like' for the other. But this way of carving up our understanding invites difficult questions: Can I truly be said to understand that you are thinking of me if I don't know what thinking of me feels like for you? Can I truly be said to understand that you are sad if I do not have an idea of how your sadness feels 'from the inside'? Or, to take a more extreme case: can I truly be said to understand that you are sad even if I myself (through some happy stroke of luck) have never once been sad myself?

I cannot begin to answer these questions here. I will say, though, that Smith generally takes understanding another's thoughts and feelings to involve either 1) drawing upon old knowledge about what it is like to be sad, or curious, or desirous of fame, and so forth, 2) producing a new and vivid idea of what the episodes of another person's life must be like for them, or both. For him, and also for those contemporary philosophers of mind who take inspiration from him, understanding another person's inner life critically involves acquaintance with that person's experience.² So, my discussion of the relation between care and understanding will naturally focus on a conception of understanding that emphasizes the importance of experience.

² Philosophers of mind who adopt this position, and who also claim to draw upon Smith, include Gordon (1995) and Goldman (1992), along with authors who have a more pronounced cognitive-scientific orientation, such as Kiesling (2012).

ii. empathy as a ‘bootstrapping’ device: the official theory

On the very first page of *The Theory of Moral Sentiments*, Smith both gives us his version of the picture of egocentric primacy and identifies empathy (which he calls “sympathy”) as the means by which get outside of ourselves, and arrive at understanding of and care for others. His identification of empathy as *the* means by which we come to understand others is straightforward: “As we have no immediate experience of what other men feel, we can form no idea of the manner in which they are affected, but by conceiving what we ourselves should feel in the like situation” (I.i.1.1). Clearly, Smith thinks our apprehension of others’ experience can only come through a mental operation that exploits our familiarity with the only inner life we directly apprehend, namely our own. His affirmation of the other half of the picture of egocentric primacy requires a little more work to see clearly, coming as it does in the first, rather infamously garbled sentences of the text. Smith writes:

However selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity and compassion, the emotion which we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. (TMS, I.i.1.2)

To see what is going on in these lines, we need to understand what Smith means by “pity or compassion.” We tend to associate ‘pity’ and ‘compassion’ with feeling bad about someone’s unhappy situation, and feeling motivated to ameliorate it. These certainly seem to be phenomena distinct from what I have been calling empathy, which involves feeling (or imagining feeling) something like what someone else feels. However, Smith’s later uses of ‘compassion’ suggest that he may not have in mind a notion of compassion that necessarily involves benevolence (see, e.g. TMS, I.i.1.10 and VI.iii.15). Rather, it is likely that he has in mind the more archaic (now obscure) sense

of compassion that just denotes participation in someone else's suffering.³ If that is correct, then we can understand Smith as holding the view that at least compassion (and possibly pity also) are a subset of the feelings generated through the process of empathizing with another. This interpretation is supported by Smith's later statement: "Pity and compassion are words appropriated to signify our fellow-feeling with the sorrow of others. Sympathy, though its meaning was, perhaps, originally the same, may now, however, without much impropriety, be made use of to denote our fellow-feeling with any passion whatever" (TMS, I.i.3.5).

Smith writes that there are "some principles in [man's] nature, which interest him in the fortunes of others, and render their happiness necessary to him," and that pity or compassion is "of this kind." I take him to be using the phrase "of this kind" to signal that pity or compassion is one of the principles in human nature that has the effect of generating both interest in other's fortunes and care for their happiness. Since Smith also describes pity or compassion as an emotion, and since an emotion does not seem to be the kind of thing that can be a principle, we can assume that what Smith really means is that our *tendency* to feel pity or compassion is one of these principles of our nature. At any rate, he seems to be endorsing the following general order of explanation: our tendency to feel pity or compassion (by which, I have argued, he means our tendency to empathize), is a principle of our nature that causes our basic care for other people's feelings and fortunes, and not the other way around.

So, according to the opening page of *The Theory of Moral Sentiments*, both understanding and basic care are to come from empathy. This is what I will refer to as Smith's 'official theory' of empathy. In a moment, I will move on to the problem that this dual role generates, but first we need to look more closely at the details of Smith's conception of empathy.

iii. empathy as imaginative projection

Smith's description of this mental operation uses various spatial metaphors that help to give a sense of what he has in mind: frequently, we are told that empathy involves "enter[ing] into" another's situation.⁴ Smith also sometimes speaks of

³ This meaning of 'compassion,' now deemed "obscure" by the *Oxford English Dictionary*, first appeared in English in the 1340's (The OED Online entry for 'compassion, n.' June 2012, Oxford University Press).

⁴ The phrase recurs a total of 73 times in *The Theory of Moral Sentiments*, including e.g. at I.i.1.4; I.i.1.8; I.i.2.2; and I.i.2.6.

“bringing the case home” to oneself.⁵ Following the lead of these spatial descriptions, I will call the operation in question ‘imaginative projection.’⁶ Properly engaging in imaginative projection means mentally conjuring up as much of the detail of another’s circumstances as possible, as though one were in the other’s position: “the spectator must, first of all, endeavor, as much as he can, to put himself in the situation of the other.” (TMS, I.i.4.6).

What, exactly, counts as another’s situation is a difficult question. Should it include only a person’s material circumstances? How about their previous experiences, or even their likes and dislikes? Smith wavers on this point. About empathizing with the mother of a dead son, he writes: “I do not consider what I, a person of such a character and profession, should suffer, if I had a son, and that son was unfortunately to die: but I consider what I should suffer if I was really you, and I not only change circumstances with you, but I change persons and characters” (TMS, VII.iii.1.4). Here, Smith is conceiving of agents’ situations in what we can call a broad sense, one that encompasses facts about our inner life, as well as our material circumstances. However, he also writes: “Sympathy...does not arise so much from the view of the passion, as it does from that of the situation which excites it. We sometimes feel for another, a passion of which he himself seems to altogether incapable, because, when we put ourselves in his case, that passion arises in our breast from the imagination, though it does not from his in reality” (TMS, I.i.1.10). This suggests that Smith is thinking of the situation with which we imaginatively engage as something more like the sum total of the other person’s material circumstances, and less like the sum total of their material circumstances, *plus* their history, their preferences, their fears, and so on. Overall, though, it makes sense to think that insofar as our goal is to understand someone, the more complete we can make our projection into her circumstances (broadly understood), the better. To illustrate: I will better understand how a lost child feels if I am able to imagine not just how I would feel if I, with my current education, size, emotional maturity, and so forth were lost, but rather how I would feel if I could not read, did not know how to ask for help, and so forth.

⁵ See e.g. I.i.1.4; I.i.2.6; I.i.3.9; and II.i.3.2.

⁶ Here I am following Charles Griswold, who describes Smith’s notion of imaginative engagement as “projective imagination” (1999: 90).

When we imagine ourselves in another's situation, Smith thinks, we actually *experience* something like a weak version of the feelings we imagine the other to be having. Smith calls this kind of feeling a "shadow": "Every man feels his own pleasures and pains more sensibly than those of other people. The former are the original sensations, the latter the reflected or sympathetic images of those sensations. The former may be said to be the substance; the latter the shadow" (TMS, VI.ii.1.2). The comparison to a shadow helps to capture the quality of the feeling attained through projective imagination: just as a shadow mimics the outer form of an object exactly, but in a duller tone, an observer's empathetic feeling has the same object and the same phenomenal quality as the original emotion, except that it is less vivid. It is this less vivid feeling that is, according to the opening of *The Theory of Moral Sentiments*, supposed to be the source of our basic care for others. The feeling is, apparently, enough to make us perceive others' fortunes as worth caring about for their own sake. In particular, it makes us see others' happiness as "necessary to us" (TMS, I.i.1.1).

For Smith, the imaginative transportation into the other's situation is typically effortful. It is true that we have a natural capacity for empathy, but we are all affected by various motivational and imaginative limitations that affect our ability to empathize, and so we often "find it...difficult to sympathize entirely, and keep perfect time" with other people's feelings (I.iii.1.8). Especially when it comes to others' sorrows, we must battle against a "dull sensibility to the afflictions of others" in order to see things from their points of view (I.iii.1.13). The task of imaginatively abstracting from our own position is, Smith thinks, hard enough that we reserve some of our highest admiration for the person who is able to feel nearly as much for others as he does for himself (TMS, VI.iii.1.1 ff.).

iv. the problem with the official theory

The idea that a weak but actual feeling, rather than the mere conception of a feeling, is needed in order for us to have basic care for other's feelings makes a certain sense given that Smith is, after all, a sentimentalist. Though his theory of sympathy is ultimately quite different from Hume's, he has no quarrel with Hume's general insistence upon the idea that only sentiment, and not reason, can ultimately motivate

us.⁷ However, the idea that we need to feel a shadow of another's feeling in order to care for him runs into a serious problem. In brief, the difficulty is this: it is hard to see how the experience of a feeling could be anything but the experience of one's *own* feeling, and the experience of one's own feeling is not a suitable source for basic care for others.

To bring out the problem more clearly, let us consider a case in which I am to empathize with another person's suffering. Suppose I have turned my attention to Jan, whose investments in the tulip market have fizzled spectacularly. Taking stock of Jan's circumstances and outward behavior, I attempt to imaginatively enter into his position. I imagine that I once was rich, but am suddenly poor, and all because of some bad bets. Let us suppose (not I think, implausibly) that part of Jan's inner state is a painful regret occasioned by the enduring presence of the thought: "I have lost my whole fortune." If I am to successfully understand Jan's inner state, I should, according to Smith, experience a shadow of this painful regret. What, precisely, will be the nature of that shadow? To answer that question, we need to think a bit about what work the indexical 'I' and the indexical possessive pronoun "my" are doing in the thought that is the focus of Jan's regret. What does Jan regret, exactly? In some sense, it would be correct to say that Jan regrets the loss of Jan's fortune. But unless he is highly unusual, it is not the loss of Jan's fortune qua *Jan's* fortune that matters to Jan. It is, rather, the loss of *his* fortune that he cares about. Jan's thought, "I have lost my fortune, " is not really interchangeable with the thought "Jan has lost his fortune," (or even "I have lost Jan's fortune"). A counterfactual will bring out the point: if by some odd fluke Jan had forgotten that he was that man whom people call Jan, then the thought "I have lost Jan's fortune" would have a resonance very different from that of the thought "I have lost my fortune."⁸ It might prompt Jan to feel guilty for having ruined another person's future, but it would not cause him to long for his bygone salad days. Clearly, Jan's regret is not for the fortune of the man named Jan, who happens to be him. Rather, it is essential to the nature of his regretful thought that the one who thinks it is the same person who has lost the fortune, and knows that he is that person.

⁷ See the *Treatise*, II.iii.iii and following.

⁸ John Perry introduced the use of counterfactual analysis to show that indexicals are essential: swapping third-personal designators in for first-personal designators inevitably alters the character of attitudes. See Perry (1993: 3-53).

The character of Jan's thought will not be accurately preserved if my shadow of Jan's regret includes the thought "Jan has lost his fortune," or "I have lost Jan's fortune." Jan's thought "I have lost my fortune" refers to the person who is thinking the thought as the one who has lost a fortune, and so my recreation of Jan's regret must have this same structural feature. The thought's indexicals must be preserved 'as is' in my shadowy imaginative reconstruction of Jan's inner life. That means that when I come to understand Jan's regret through imaginative projection into his situation, I myself will be pained (if only a little) by the thought, "I have lost my fortune." Otherwise, my experience of the putative shadow of his inner life will not be accurate. But in that case, the only care that painful thought should generate is care for my own fortune, and not care for Jan or for his fortune!

This conclusion needs a minor qualification: my regretful feeling could cause me to be concerned about Jan's regret, but only instrumentally. If empathy with painful sentiments will itself be a bit painful, and if I am bound to empathize at least some of the time, it would be natural for me to prefer that others not have such sentiments, and so I might hope for the return of the man's fortune. However, this is not basic care, as I have defined it, and it is presumably not the kind of concern for other people that Smith has in mind. After all, my preference in this case would be just as well served by any circumstance that interrupts the empathetic mechanism, including my falling asleep or simply forgetting about Jan. Such purely instrumental concern for others' feelings is still fundamentally selfish.

Effectively, Smith is caught in a bind. According to his picture of how empathy works, we understand a person's feeling insofar as we are able to accurately reproduce the other's passion in us through imaginative effort. When we succeed in understanding the other, only a variation in liveliness distinguishes their passion from the shadow of it that we experience. However, the development of understanding looks to be at odds with the generation of care. My empathetic experience of someone else's regret about their situation can only result in my feeling regret about my own situation—a particularly strange result, given that I may well be in an enviable situation, myself.

A confused grasp of Jan's regret does look like it could generate care for Jan, of a sort. If my imaginative reconstruction of Jan's feelings included a bitter regret focused on the thought "*Jan* has lost his fortune," then my empathetic feeling would involve a

shadow of regret directed at Jan's loss of his fortune, rather than my loss of my fortune. If I come to regret Jan's having lost his fortune, then perhaps it makes sense that I would come to care about Jan and his situation (although I am not sure about that). But getting to care in this way means sacrificing a real understanding of Jan's inner life. Jan's regret does *not* focus on the thought "Jan has lost his fortune," and if my reconstruction of Jan's regret includes this kind of focus, then I have misunderstood Jan's inner life in a significant way. Furthermore, there is something suspect about the kind of care we could expect this kind of reconstruction to produce. Perhaps it would not seem so odd for me to care about Jan's loss in the way that a regretful feeling would demand. But suppose that the feeling in question were François' unhealthy self-hatred. It would be utterly bizarre, and not at all consistent with our expectations of what basic care should amount to, for my care for François to arise from empathetic self-hatred (or, depending upon how we think of François' feelings, from hatred of François!).

Schopenhauer described the question of how our sharing others' feelings could lead us to care for others as "the great mystery of Ethics, its original phenomenon, and the boundary stone, past which only transcendental speculation may dare to take a step" (170). Smith's official attempt to dispel this mystery, whilst also accounting for our basic understanding of others' passions, does not succeed. His account draws on what strike us as familiar thoughts: it seems reasonable enough to think that I understand you at least in part by relying upon my ideas of my own experience, and that my feeling something like what you are feeling has something to do with my caring for you (witness the prevalence of injunctions to "walk a mile in my/their shoes"). However, when we try to determine exactly how our empathy with others relates to our understanding and care for them, difficulties like Smith's rapidly crop up.

Later on in the *The Theory of Moral Sentiments*, Smith does qualify his 'substance : shadow :: original feeling : empathetic feeling' analogy in a way that indicates he sees a problem with his original take on the relations between empathy, care, and understanding. Of those observed by another, Smith writes: "What they feel will, indeed, always be, in some respects different from what he feels...because the secret consciousness that the change of situations, from which the sympathetic sentiment arises, is but imaginary, not only lowers it in degree, but, in some measure, varies it in kind, and gives it a quite different modification" (TMS, I.i.4.7). Smith does not say what

this “quite different modification” is supposed to be, and he does not mention anything like it again. Still, his appeal to it may signal a kind of recognition that an exact copy of another’s emotion, taken by itself, is not a realistic source of care for the other. Perhaps without realizing it, Smith seems to be grasping for a way of having his cake and eating it too. That is, he is trying to ensure that the empathetic emotion is different enough from the original that it can be a source of care for the other, whilst also holding on to the idea that the one is enough like the other that it could be a source of true understanding of the other’s inner life. The fact that Smith is only willing to talk about the hypothetical modification in the most vague of terms does not bode well for this strategy, however.

In the next and final section of this paper, I am going to turn from criticism to an attempt at a constructive modification. I will argue that Smith should abandon his commitment to the idea that empathy generates care for others. Doing so will dissolve the problem I have discussed above, but this is not the only reason to reject the idea that empathy generates care. Whether he realized it or not, Smith’s account of our motivation for empathizing with others practically entails that basic care is at least sometimes a precondition of empathetic imagination. So, giving up on his particular strategy for countering psychological egoism will make his theory more coherent in two respects.

III. Care and Understanding: the Implicit Smithian Alternative

i. sentimental harmony and the motivation to empathize

As I have mentioned, Smith describes imaginative projection primarily in terms that make it look like it involves deliberate work. For Smith, perfect fellow feeling is a goal that we struggle to meet. He writes of the spectator’s task: “the spectator must, first of all, endeavor, as much as he can, to put himself in the situation of the other, and to bring home to himself every little circumstance of distress which can possibly occur to the sufferer. He must...strive to render as perfect as possible, that imaginary change of situation upon which his sympathy is founded” (TMS, I.i.4.6). Smith’s talk of “endeavoring” and “striving” makes it clear that this is an difficult undertaking; it may be natural, but it is hardly automatic. Because Smith rejects the conception of imaginative projection as automatic or effortless, he is faced with this question: why are

we motivated to perform the imaginative exercises that generate fellow-feeling in the first place? In what follows, I will argue that Smith ought to treat basic care as a precondition of a kind of imaginative projection that is central to human sociability, the only kind that Smith pays any real attention to. Were it not for the attitude of basic care, we would not be driven to engage in imaginative projection in a large and important set of cases. As we have seen, for Smith, imaginative projection is the essential means by which we acquire experiential understanding of others. Taking the two previous points together, we will be able to conclude that in a large and important set of cases, we would not come to understand the target agent if we did not already care about him.

We have no particular reason to assume that there is just one explanation for why we engage in imaginative projection. It seems obvious that imaginative projection could sometimes be helpful for securing benefits for, and avoiding harm to, ourselves. For instance, if I want to avoid losing my lunch money, it could be helpful to project myself into my bully's position, and imagine just what kind of groveling I would find most appealing if I were him. But Smith is quite clear that we do not only engage in imaginative projection when knowing a person's mind could be important to our fortunes. It is hard to see how fellow-feeling with another person could be useful to us in cases where the other is unlikely to ever interact with us, and yet Smith is adamant that our fellow-feeling can and does extend to such people: we are liable to "bring home to ourselves" the misery of "any innocent and sensible being," not just the misery of those in our own country (TMS, VI.ii.3.1). So, in cases where knowing the other's mind does not seem useful to us, why do we nevertheless engage in imaginative projection?

Smith does have something to say about why we are motivated to engage in imaginative projection. And reflection upon the explicit reason he gives for our motivation will reveal that Smith should not endorse the claim that empathy produces caring. According to Smith, we "passionately desire" to attain a "harmony of hearts": we want our feelings to match others' (TMS, I.i.4.7). This desire is at work in our attempts to reconcile our opinions about things like art or philosophy with the opinions of others. As Smith points out, sharing others' opinions in these cases does not require "sympathy, or that imaginary change of situations from which it arises, in order to produce, with regard to these, the most perfect harmony of sentiments and affections" (TMS, I.i.4.2). But Smith argues that the desire for harmony is still stronger in cases

where attaining a harmony of hearts means entering into another's position, and understanding things from that person's point of view. When we cannot harmonize our feelings with those that an agent feels about his surroundings, the result is "shock" and "pain" (TMS, I.i.2.1).

Smith identifies the harmony of hearts as the source of great pleasure for us: "[N]othing pleases us more to than to observe in other men a fellow-feeling with all of the emotions of our own breast" (TMS, I.i.2.1). When we fail in our efforts to bring our own feelings in line with those of another person, conversely, the result is a kind of pain: "it hurts us to find we cannot share his uneasiness" (TMS, I.i.2.6). Smith is careful to clarify, in response to an objection from Hume, that even when we feel another's pain, which is at least partly an unpleasant experience, the fact of our feelings' correspondence is itself an unfailing source of happiness.⁹ Of the silence and mirth of our companions, when we ourselves are mirthful, Smith writes: "this correspondence of the sentiments of others with our own appears to be a cause of pleasure, and the want of it a cause of pain, which cannot be accounted for" except by the thought that the correspondence is an independent source of pleasure (TMS, I.i.2.2). Note that Smith is not arguing that we find sentimental harmony pleasurable because it is good for us in some other way. We can of course imagine cases where this would be true: I am pleased that we both enjoy scary movies, because it means I don't have to watch *Alien* alone. Or, referring back to the case of the bully, feeling something like what the bully feels could be pleasurable, because I take pleasure in having the resources to avoid the bully's extortion. But on Smith's view there is something pleasurable about the sharing of sentiment regardless of any consequences. Sentimental harmony, it seems, is pleasurable in and of itself.

Smith does not pin down the relationship between the pleasurableness of sentimental harmony and its desirability as precisely as we might like. Pleasure makes its first appearance in the opening lines of *The Theory of Moral Sentiments*: "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it" (TMS, I.i.1). We might wonder, here, what explanatory role (if any) pleasure is playing: should we think

⁹ See Hume's letter to Smith dated 28 July 1759 (letter 36), in Smith (1987).

of the “principles” as directly motivating our interest in and care for others, whilst also making it the case that we take pleasure in sharing others’ pleasures? Or should we instead think that the “principles” are the facts of our nature that make it the case that sentimental harmony is pleasurable for us, and that the prospect of pleasure is in fact what directly motivates us to empathize? In other words: do we want sentimental harmony because it feels good, or is the fact that it feels good just an indication of its inherent desirability? Smith’s attempts to characterize our desire for sentimental harmony do frequently invoke pleasure. Furthermore, Smith explicitly adopts the position that “Pleasure and pain are the great objects of desire and aversion” (TMS, VII.iii.2). These considerations encourage the thought that for Smith, it is the *pleasurableness* of sentimental harmony that directly motivates us to pursue it.¹⁰ However, there is simply not enough textual evidence to draw that conclusion with any real confidence, and one could argue that reading Smith in that way makes it look as though he accords pleasure an unrealistically prominent role in the explanation of empathy. For our purposes, however, this interpretive question is not critically important. The important point is that Smith regards sentimental harmony as an end that is either inherently desirable and thus capable of directly motivating us to empathize, or as an end that is desirable because it is pleasurable.

Before moving ahead, let us sum up what we have so far established. According to Smith, all experiential understanding is the product of (successful) imaginative projection. Smithian imaginative projection is typically effortful, so it makes sense to ask why we are motivated to engage in it. One reason to engage in it is that it provides an understanding of others that could be useful to ourselves. However, we have reason to think that not all imaginative projection is motivated by such benefits. Smith himself emphasizes the fact that we are motivated by a desire for sentimental harmony, which is itself pleasurable. Now, why do people find sentimental harmony desirable and pleasurable, even in cases where there is no instrumental benefit to be accrued from the harmony? Smith does not give us an explicit answer. My purpose here is just to show that on Smith’s view, basic care must precede the experiential understanding empathy provides in a large and important number of cases. Therefore, I am not going to attempt to give a full account of why we find sentimental harmony desirable and

¹⁰ Or, alternatively, we might say that it is the pain of disharmony that drives us to empathize with others.

pleasurable. Rather, I am going to argue that we can only make sense of the distinct pleasure that we take in sentimental harmony in and of itself, a pleasure that in turn appears to motivate our imaginative projection, if we conceive of ourselves as feeling basic care for others *prior to* imaginative projection.

ii. care's role in the motivation to empathize

What would our attitude toward others have to be like in order for us to take pleasure in sharing their feelings, in cases where no instrumental benefits are at stake? At this point, I see two possibilities. The first is that there is nothing in particular that our attitude toward others would have to be like. On this view, the pleasure we take in the correspondence of sentiments would not in any way be contingent upon our regarding others' sentiments as independently valuable or worth knowing about. The other possibility is, of course, that the pleasure we take in the correspondence of sentiments *is* contingent upon our regarding others' sentiments as independently valuable or worth knowing about. How can we decide between these two possibilities? Smith is all for psychological realism. Therefore, I propose that we take a cue from him, and begin by considering whether the first view seems like a good fit with our pre-theoretical familiarity with human motivation.

One might argue that harmony *just is* a pleasing state to be in, regardless of whether we care about others' opinions as such. It just feels good to have the same feelings that someone else does about the new blockbuster, or the politician's behavior, or the German football team. It may not be immediately obvious that sharing sentiments couldn't feel pleasurable in itself in the absence of a pre-existing basic interest in others' feelings. Upon reflection, however, it seems hard to imagine how it would be psychologically possible to be motivated to seek sentimental harmony just because it is inherently pleasurable, without also being interested in or concerned with others' feelings.¹¹ It would be something like being motivated to keep a promise by the thought that breaking a promise would make one feel bad, even if one does not care

¹¹ Dick Moran has suggested that a 'negative' version of this possibility might be more psychologically plausible. That is, it might be psychologically accurate to claim that we could find sentimental *disharmony* *painful* even if we did not have pre-existing care for others. I am unsure what to say about this suggestion. Personally, this possibility does not strike me as any more realistic than did the corresponding claim about the pleasurableness of harmony, but I recognize that my intuitions on this point may be idiosyncratic.

about promises themselves. Why would one feel bad about breaking the promise if one did not care about promises as such?

It is easy enough to think of cases in which we might value a correspondence of our attitudes with those of others without considering others' attitudes to be independently valuable or interesting. However, in every case I have been able to conjure up where we take pleasure in correspondence without considering others' attitudes to be independently valuable or interesting, the pleasurableness of the correspondence seems to be due to the instrumental benefits it affords. Take this example: I am pleased to share your positive feelings about the German football team. Now, it seems possible that I could be pleased by this without having any antecedent interest in your opinions. It could be that I just want the cheers for the German side to be loud, in order to increase their chances of winning. But suppose there are no instrumental benefits to our correspondence of sentiments in this case. Could I still desire that our sentiments correspond, without being interested in or concerned with your opinion as such? Such a desire just might be psychologically possible, but it strikes me as improbable and strange. It seems to me that the pleasure I take in our shared feelings about the German team must either be covertly instrumental, or dependent upon my pre-existing interest in others' attitudes. And in some cases, even a pleasure in sentimental harmony that *is* instrumental will nevertheless be dependent upon a pre-existing care for others' sentiments as such. For instance, I may take pleasure in our having the same sentiments about the German team at least in part because it means I will have someone to cheer along with. But it could very well be that I would not enjoy cheering along with you unless I regarded your sentiments as independently interesting, something to be cared about.

I have argued that since Smith treats imaginative projection as effortful, it makes sense to ask why we are motivated to engage in it. On occasion, the instrumental benefits to be accrued from sentimental harmony, and the pleasure derived from those instrumental benefits, may motivate imaginative projection. However, Smith does not discuss such cases. He is interested in cases where we are motivated to imaginatively project just by the prospective pleasurableness of sentimental harmony itself. And when it comes to these cases, an attitude of interest or concern directed at the other— that is, an attitude of care— seems to be a precondition of taking pleasure in sentimental

harmony. Since we are talking about cases in which instrumental interests are not at play, we can further specify that the care in question is basic care. To be sure: this line of reasoning does not amount to a deduction that basic care is sometimes a precondition of imaginative projection. There is no abstract, conceptual consideration that dictates this conclusion. Rather, the conclusion's warrant is based upon an observation about human psychology, one that is admittedly not drawn from a complete and thorough survey of our motivational profiles. Still, it seems to me that it hits on a true fact about the ways we take interest in other people's lives.

If Smith were to conceive of the relation between basic care and imaginative projection (and, by extension, the relation between basic care and understanding) in the way that I suggest he should, given his explanation of our motivation to empathize, then he would not in fact end up in the bind in which his official theory is entangled. It may already be obvious why the alternative conception of the relation between care and understanding avoids this problem, but let us make it explicit by employing one of Smith's own examples of an effort to achieve sentimental harmony.

iii. an illustration of the amended theory

Late in *The Theory of Moral Sentiments*, Smith describes the operation of imaginative projection in the case where one is relating to a widow or widower:

When I sympathize with your sorrow or indignation, it may be pretended, indeed, that my emotion is founded in self-love, because it arises from bringing your case home to myself, from putting myself in your situation, and thence conceiving what I should feel in the like circumstances. But though sympathy is very properly said to arise from an imaginary change of situations with the person principally concerned, yet this imaginary change is not supposed to happen to me in my own person and character, but in that of the person with whom I sympathize. When I condole with you for the loss of your only son, in order to enter into your grief I do not consider what I, a person of such a character and profession, should suffer, if I had a son, and that son was unfortunately to die: but I consider what I should suffer if I was really you, and I not only change circumstances with you, but I change persons and characters. My grief, therefore, is entirely upon your account, and not the least upon my own. (TMS, VII.iii.1.4)

Now, let us consider the example of sympathetic grief in the light of what we have said about the relation between care and understanding in Smith. In an effort to attain an understanding of your grief, to reach sentimental harmony with you, I put myself in your position by imagining being you. If the effort is successful, I will imagine feeling what you are feeling. I will feel devastated "as" you. I will feel grief at having lost

my child. This feeling could in fact be described as selfish or self-directed, in the sense that it concerns me and what was in some sense mine. Smith's claim that I "enter into your grief" by imagining being "really you" suggests that this is what has in mind. However, this feeling of mine can serve an informational function in the effort to arrive at a feeling that is *not* self-directed. As I have argued, it cannot spontaneously generate basic care for you. But what if I *already* care about you? What if I am already interested in your experiences? What if your feelings matter to me, even before I have experienced an echo of them? In that case, it makes sense that I would be motivated to put the feelings I derive from imaginative projection to use. In particular, I can infer that you are experiencing something like the grief that I am experiencing (faintly) in virtue of my imaginative engagement with your situation. And now that I have some idea of what your experience must be like, even though I cannot directly experience it myself, my general care for you is better informed. This will help me to demonstrate my understanding of your plight and otherwise condole you.

IV. Conclusion

From Smith's remarks about our desire for sentimental harmony, I have drawn out an implication that is in tension with some of Smith's more explicit pronouncements concerning the relation between care, empathy, and understanding. Smith officially identifies the empathetic mechanism as the source of both our interest in others and our concern for them. However, Smith also maintains that we engage in imaginative projection because we desire sentimental harmony (either because it is inherently pleasurable or because it is a good in and of itself). And reflection upon our own psychological make-up supports the conclusion that an orientation of basic care is in fact a condition of finding sentimental harmony non-instrumentally desirable. Therefore, if Smith 1) wishes his theory to be coherent and 2) does not want to give up his account of why we are motivated to empathize, he should treat basic care not as the product of imaginative projection, but as a something that often precedes it. In many cases, basic care might even be regarded as the driving force behind the imaginative effort in question.

As we have seen, Smith's official account faced a major problem. The two tasks he assigns to the empathetic mechanism are at odds. For Smith, the key to understanding another is to arrive at conception of their passion that is as close as possible to the original, an achievement possible thanks to the mind's imaginative capacity. This same mental process is also supposed to generate care for the other. As the case of Jan's regret made clear, a truly accurate 'shadow' of another person's feeling is not the kind of thing that can generate the right kind of care for the other. It might be misleading to say that the approach I have suggested on Smith's behalf would solve this problem. It would be more apt to say instead that it *dissolves* the problem. In effect, it prevents the problem of accounting for both care and understanding through empathy from arising, by scaling back on the explanatory ambitions for empathy.

As I hinted above, empathy could be relevant to how we direct and manifest our basic care. I might be generally interested in Jan, and think of his inner life as mattering to me, but it is only when I come have a sense of his regretful feelings that my care is properly informed. If my care is of the good-willed variety, I might use my understanding to offer Jan support; I might, for instance, let him know that I once suffered from similar feelings of regret, but was able to overcome them by refocusing my energies on work (other than tulip speculation!). That said, if Smith were to opt for the approach to care, understanding, and empathy that I have claimed would be consistent with his story about why we imaginatively project, then he would have to give up on the idea that basic care actually emerges out of empathy.

In the first section of this paper, I suggested that Smith's response to the twin skeptical threats of psychological egoism and solipsism is shaped by his commitment to the picture of egocentric primacy. This picture, which treats our own self-knowledge and self-care as uniquely primitive and not in need of explanation, has a *prima facie* plausibility that is reflected in its considerable philosophical staying power. Many of Smith's philosophical descendants endorse this picture to some degree. Indeed, the current wave of interest in empathy seems to be largely propelled by the thought that we somehow, as Nancy Sherman puts it, "step beyond the egocentric point of view" (1998: 83). Claims like Sherman's take it for granted that we begin with self-understanding and self-care, and must somehow turn outwards in order to relate to others as minded beings whose inner lives matter. If Smith were to adopt the view that

basic care precedes empathy, as I have suggested he should, would he need to rescind his commitment to the perennially popular picture of egocentric primacy?

His commitment to the epistemic component of the picture of egocentric primacy, the thought that we cannot feel others' feelings directly, need not be affected. For better or for worse, the improvement upon Smith's explicit account that I have proposed retains the assumption that our understanding of others' minds must come though bootstrapping up from our primitive self-understanding. When it comes to the other half of the picture, the story is more complicated. I have argued that Smith should recognize basic care as a necessary precondition for much of our empathetic engagement with others, and for the understanding that this engagement produces. Strictly speaking, this recognition would not undermine the claim that our care for ourselves is primitive and needs no explaining, whereas our care for others is not similarly primitive and does need to be explained. It does undermine Smith's official explanation of our basic care for others, however. And so we are left with two possibilities: Smith would either need to supply a new explanation of why we feel basic care for others, one that does not rely on empathy to move from care for ourselves to care for others, or he would need to forego the search for such an explanation and accept basic care for others as a primitive psychological disposition on a par with our self-care. With empathy eliminated as an explanation for basic care, though, it is very difficult to see what realistic alternative explanations might be available to Smith if he were to pursue the former option. Therefore, although the claim that basic care precedes empathy does not strictly require that we abandon the care component of the picture of egocentric primacy, it does place would-be defenders of the care component of the picture in a difficult position.

This paper has been focused on the particular ways that Smith appeals to empathetic mechanisms in his theory of mind and morals, but the problems that Smith encounters are not just the result of minor quirks in his descriptions of human psychology. His problems begin with the picture of egocentric primacy. When we posit such a dramatic gap between our how we relate to ourselves and how we relate to others, it is entirely natural that we should attempt to bridge that gap by appeal to self-care and self-understanding, phenomena with which we are so intimate that they seem to require no explanation. However, empathy-based solutions to the problem of

accounting for care for others run into the problem of appropriately acknowledging the distinction between self and other. Reproducing something like others' feelings 'in' ourselves does not seem to actually get us to care for the other. And, while I did not discuss the question in this paper, we might also ask whether this kind of reproduction actually gets us to a real grasp of others' passions, or if it only gets us to an idea of our own passions that we mistake for an idea of others' passions. Ultimately, I think that these and related problems give us a reason not just to look for alternative ways of bridging the gaps between self and other, but to look at the picture of egocentric primacy, which posits these gaps in the first place, with a specially critical eye.