8

Empathy and Empirical Psychology

A Critique of Shaun Nichols's Neo-Sentimentalism

Lawrence Blum

Experimental philosophy often characterizes and belittles traditional philosophical inquiry as 'armchair' and counterposes it to going out in the world to see how things actually operate by running experiments or at least drawing on experiments run by others. Shaun Nichols is a leading proponent of the experimental philosophy method. His book *Sentimental Rules: On the Natural Foundations of Moral Judgment* (Nichols (2004)) makes extensive use of psychology experiments and findings in the course of building an argument about the nature of and close connection between emotion and morality.

We can roughly distinguish two traditions in philosophical thinking that assert a close relation between emotion and morality, each rejecting a different form of rationalism. The first, which has come to be called 'sentimentalism', sees emotion, rather than reason, as the source of moral *judgment*. This tradition originated with the Scottish moralists, especially Hume and Hutcheson, and has had something of a revival in contemporary ethics (e.g. Blackburn (1998); Gibbard (1990); D'Arms & Jacobson (1994)). Nichols locates himself in this tradition, neo-sentimentalism, while staking out a distinctive position within it (discussed below). This view rejects a rationalism that takes the form of privileging reason as the foundation of (either as source, or as a mode of access to) moral judgment.

The second tradition sees emotion as required for the human capacities necessary for leading a moral life, or being a moral person—that is, required for moral motivation or moral responsiveness more generally. Schopenhauer is an important exponent of this view. In his *On the Basis of Morality* Schopenhauer criticized Kant's view that reason or rationality is fundamental to moral capacity—that reason alone (or predominantly) provides the motive for moral action—and suggested instead that the emotion-based motive of compassion was fundamental (Schopenhauer ([1841]/1965)). This tradition of concern with motive, action, and response is less concerned with moral judgment. Nichols does not explore or acknowledge this second tradition of thinking about the relation between emotions and morality; but parts of his argument bear on it.

Nichols agrees with Hume that moral judgment rests ultimately on sentiments or emotion rather than reason. He says that 'core moral judgment' involves an affective mechanism of response to distress, as well as a system of rules that he refers to as a 'normative theory'. Core moral judgment involves distinguishing the conventional from the moral—using one's fork to eat vs. not harming others. The ability to make this distinction shows up quite early in children. Nichols argues that the distinctive force and status of moral rules (among rules in general) is dependent on their eliciting strong affect, and this can happen only in persons who are able to respond emotionally to the suffering of others. For example, psychopaths are defective in their ability to make moral judgments, although they recognize the social existence of norms that forbid causing harm, and they also fail to respond emotionally to suffering in others. Nichols suggests that the emotional deficiency is the source of the failure to appreciate the moral character of the norms.

I will argue that Nichols's view suffers from several deficiencies: (1) It operates with an impoverished view of the altruistic emotions (empathy, sympathy, concern, compassion, etc.) as mere short-term, affective states of mind, lacking any essential connection to intentionality, perception, cognition, and expressiveness. (2) He fails to keep in focus the moral distinction between two very different kinds of emotional response to the distress and suffering of others—other-directed, altruistic emotions that have moral value, and self-directed emotional responses, such as personal distress, that do not. (3) Nichols is correct to see morality as requiring affectivity, the capability of emotional response to others; but his incorrect view of altruistic emotions (and of emotions in general) leads him to misstate the connection between morality and emotion. (4) Nichols fails to recognize Schopenhauer's form of anti-rationalism as distinct from Humean sentimentalism; some of his arguments presented to support the latter instead lend support to the former. (5) Finally, while agreeing that moral philosophy is strengthened by knowledge of empirical psychology, I suggest that the foregoing failures of Nichols's argument are partly due to his misuse of particular empirical results and findings, his being over-enamored of empirical psychology, and possibly in part to a weakened commitment to the distinctive contribution that the humanistic methods of philosophy make to our understanding of the moral enterprise.

8.1 Nichols on Empathy as Having the Same Feeling as the Other

In the second chapter of *Sentimental Rules*, 'Sparks of Beneficence: The Varied Emotional Responses to Suffering in Others', Nichols turns his attention to the affective system of response to suffering for the emotion-based component of his sentimentalism. He also says he is interested in the underpinnings of altruistic motivation, presumably because altruistic motivation is a core dimension of morality. 'My goal is restricted to the project of determining the cognitive and affective mechanisms underlying basic altruistic

motivation' (2004: 33). These two foci pull Nichols in two different directions in the chapter. The focus on altruistic motivation as a distinctive type of motivation seems addressed to a Schopenhauer-like concern with moral motivation, on the assumption that altruistic motivation is a type of moral motivation. By contrast, the focus on emotional responses to distress blurs the distinction between moral and non-moral sentiments and emotions, since all Nichols needs for his sentimentalism is that the emotional response to distress be a strong one, not that it be distinctly altruistic, or otherwise moral, in character.

Nichols begins by mentioning empathy, regarded by many psychologists as providing the motivation for pro-social behavior, psychologists' term for behavior intentionally aimed at helping others (2004: 31). He defines empathy as '"a vicarious sharing of affect" or an emotional response in which the emotion is "congruent with the other's emotional state or situation"' (ibid.). Although it is not entirely clear whether Nichols regards this definition, which he attributes to a leading social psychologist in the empathy field (Eisenberg), as entirely satisfactory, he never proffers a different one, and appears to use this one in his subsequent argument.[1]

Nichols wants to make the point that we have a range of emotional responses to distress other than empathy, and these need to be recognized both for their own intrinsic interest, and also because of their role in providing what he sees as the emotional underpinning of moral norms against causing harm. The most important of these other responses are: 'emotional contagion' (being caused to have the same mental state as another by being in that individual's presence), and 'personal distress' (being caused to have one of a number of mental states, e.g. alarm, anxiety, upset, distress by another's distress).

We ordinarily think of empathy as a moral phenomenon, in contrast to these other responses (which Nichols sometimes calls 'self-directed'); but Nichols's definition fails to capture that dimension. Having the same affect as another, and having one appropriate to another's emotional state, are morally indifferent. If Don is sad and Juan is also sad, Juan's sadness by itself is of no moral significance (in relation to Don), nor is it if we add the condition that Juan is caused to be sad by Don's being sad. This is because without Juan's recognition of or concern for Don's sadness the connection to the other required for morality in this context is absent. Nor would merely having the same feeling, as described, normally be called 'empathy' in the usual use of that term. Empathy involves a feeling *for another*, that is, a feeling whose object is another's state of mind or situation. When Nichols discusses 'self-directed' feelings, such as personal distress, he draws that terminology from the psychologist Batson, who introduces it to contrast with other-directed emotion. Nichols appears to follow Batson's implication that this is a contrast with moral significance. A self-oriented feeling normally lacks moral significance, while feeling directed toward others may possess it.[2]

[1] 'In Eisenberg's view, pure empathy is not other-oriented' (Eisenberg (2000): 671). Eisenberg is referring to herself, and thus to her view, in the third person here.

[2] Some self-oriented feelings do have moral significance, for example, a feeling of outrage at having been denied one's rights or one's dignity having been violated. But unless it involves a moral principle,

Empathy is an instance of a larger category of emotion that one might call (following Max Scheler) 'fellow-feeling' (*Mitgefühl*) (Scheler (1954)). Compassion, sympathy, and concern are other species of fellow-feeling. In fellow-feeling, we recognize that the other has a particular state of mind or situation, and we have a regard for her well-being with respect to that state of mind or situation. But we do not, as part of fellow-feeling itself, adopt or come to have that state of mind ourselves. If my friend is disappointed that she failed to attain a professional honor that she hoped for, I may be empathetic toward her with regard to this disappointment. But in being so I am not myself disappointed; I am not in a state of disappointment. I am disappointed *for her*; that is, I am acutely aware of her disappointment and perhaps I envision her disappointment firmly in my own mind. This is entirely compatible with my feeling quite satisfied myself, feeling that my own life (in contrast to my friend's) is going very well. In sympathy, I also have some affect myself, but my state of mind is directed toward another in light of (what I take to be) something bearing on her well-being. The difference between sympathy and empathy is not entirely clear, and the common usage of these terms may not consistently mark a single distinction. Perhaps empathy implies a stronger and more acute imagining of what the other is going through than does sympathy; but neither involves actually *having* the same state of mind as the other, contrary to Nichols's definition.

8.2 Fellow-Feeling as Directed Toward the Other's Well-Being, and Some Faulty Accounts

All the different forms of fellow-feeling have the following features: (1) they involve affectivity on the part of the observer; (2) they involve having as an intentional object another person in light of her state of mind or situation (or the state of mind or situation itself);[3] and (3) the aspect of the person's state of mind or situation that is conceptually tied to the fellow-feeling is the other person's well-being.

Fellow-feeling's essential focus on the other's well-being brings with it a fourth feature, a motivational dimension. Someone who has fellow-feeling for another has some concern for the other's well-being, and thus an at least minimal *ceteris paribus* motivation to help the other. In any given situation there may be several motivational pulls, some of them running contrary to such help, so fellow-feeling does not always

self-directed feeling is morally neutral. Batson's terminology of 'other-directed' is not meant to include every feeling or emotion directed toward others rather than toward the self or the self's situation. For negative emotions such as dislike, malice, or hatred are directed toward others, yet are not morally good. What he intends by 'other-directed' emotions must involve a positive regard for the other's well-being in some way—a pleasure in the other's good fortune, a worry about impending harm to her, or a concern for her distress. Schopenhauer captured these distinctions in his threefold differentiation of fundamental motives as they bear on morality—egoism (desire of one's own weal), malice (desire of others' woe), compassion (desire of others' weal) (Schopenhauer ([1841]/1965): 145). Only compassion, according to Schopenhauer, is a moral motive, a motive with moral worth, and this seems to be Batson's view, which Nichols implies that he adopts.

[3] We might say 'I empathize with Reginald in light of his sadness' or 'I empathize with Reginald's sadness'. These seem very close in meaning, if not identical.

translate into helping. Nevertheless, if someone says that he empathizes with Rae with regard to a difficult emotional situation Rae is currently going through, yet is unwilling to proffer a minimal comfort that would cost him nothing and would seem an appropriate and helpful action for Rae, the idea of his having empathy for Rae starts to lose its grip.[4] In this sense there seems to be a conceptual connection between having fellow-feeling for another and motivation to help that other.[5]

What is the affective state involved in fellow-feeling if it is not the same as or homologous to that of the person whose state of mind or situation is its object? More precisely, what is its felt quality? What is the phenomenological quality of being 'sad for X' when one is empathetic toward X in relation to her sadness, that differs from 'being sad' itself? I am not sure how to think about this. Some accounts suggest a 'vicarious' sadness; but I am not sure if this is not simply another way of saying one is 'sad for' the other and so does not serve to explain it. Others suggest the feeling is a less intense form of the feeling the other has. Perhaps. But what I want to emphasize here is that fellow-feeling, including empathy, is both affective and intentional.[6] And yet we may not always be able to characterize the phenomenological feel of the fellow-feeling better than that. So the precise characterization of that phenomenal feel may not be central to our understanding of empathy, sympathy, and care. And that particular phenomenological feel may be relatively insignificant for the moral significance of fellow-feeling.

Moreover, although fellow-feeling is necessarily affective, it does not necessarily involve an occurrent affective state at every moment of which it is true to say that one has fellow-feeling for X. This is most obviously true in the verbal forms of fellow-feeling: 'I empathize with X', 'I sympathize with X', 'I am concerned about X'. Those characterizations can hold of me for an extended stretch of time during which I do not necessarily experience feeling-states of empathy, sympathy, or concern for the entire duration. But the same is true of 'have empathy (sympathy, concern) for X'. Though these attribute the noun form of the emotions to the subject/observer, they do not entail that the subject is in a particular affective state for the entire period of which the attribution is true. As Peter Goldie notes, there is a difference between having an emotion and having an episode of emotional experience (2001: 2). We might use the single term emotion for either one, but we would more normally express the episodic form by saying 'I am feeling sympathy (empathy, concern) for Midori at the moment', while the former is expressed as 'I have sympathy for Midori' or 'I sympathize with

[4] In many situations, there may be no motivation to help because help is impossible or the observer is simply in no position to help and realizes this from the get-go.

[5] I think 'sympathy' is sometimes used in a very minimal way to mean little more than 'acknowledge' but with the barest implication of caring or concern for the other. 'Sure, I sympathize with Ashok. It's too bad what's happening to him'—but with the implication that I am not really very concerned and am not prepared to lift a finger to help him. But normally, I think we would question the accuracy or sincerity of someone's saying that he had or felt sympathy for someone if he was unwilling or uninclined to take any steps to help when the cost to himself is negligible.

[6] Peter Goldie (2001) emphasizes this point.

Midori'. In this latter use there may in many cases be no phenomenological feel in the state of mind of the one who has fellow-feeling.

8.3 Missing the Intentional, Other-Directed Character of Empathy and Fellow-Feeling

As mentioned, in his definition of empathy, Nichols is reporting a widespread view of empathy within experimental psychology.[7] His (and Eisenberg's) view is not alone in failing to capture the distinctive, morally significant, features of fellow-feeling, and of empathy specifically. In a thoughtful analysis of empathy, the philosopher Nancy Snow says that it involves the observer (O) feeling the same feeling (S) as the victim (V), and feeling it *because* V feels it, knowing that she feels the same feeling as V.[8] In support of the observer and victim's having the same emotion, a feature I am challenging, Snow gives the following example: 'When Barbara feels the same emotion as Bob, they both feel sad about the same fact, Bob's father's death' (1991: 66). But while Barbara might feel sad about Bob's father's death, that is not part of her empathy for Bob. The latter requires only that she feel badly *for Bob* because of his sadness over his father's death. If she knew nothing of Bob's father, and thus had no basis for personal sadness at his death, she would or could still feel empathic sadness for Bob.

Other accounts of empathy also fail to capture the intentional and other-directed structure of empathy that is a necessary element of its moral character. Sober and Wilson recognize the distinction between feeling X for someone, and feeling X full stop (Sober & Wilson (1998): 234). They express the difference by saying that Lakisha's empathy for Victor entails that Victor's feeling S causes Lakisha to feel S. (Snow's account contains this feature also.) But this causal relationship does not capture the intentional one it is meant to. Lakisha might know that Victor is depressed, and his being depressed (or, Lakisha's knowing that he is depressed) causes her to be depressed. Lakisha might say that his depression 'brings her down'. (This is an example of 'emotional contagion', mentioned earlier.) Here Lakisha and Victor have the same emotion and his emotion has caused hers. But this is not her empathy for him. The two could be combined. Lakisha could genuinely have empathy for Victor in his depression, and could also come to be depressed herself because of his depression. It may not always be easy to distinguish these in practice. When I 'absorb' my friend's anxiety or sadness, as we sometimes say, this can have a quality both of emotional contagion and of empathy. But this 'absorbing' is less likely to occur with regard to other objects of empathy, such as disappointment or physical pain. In any case, the contagion and the empathy remain two distinct psychic phenomena.

[7] Here is a similar definition from a recent article in psychology: 'empathy (i.e. a vicarious emotion that results in feeling the same as another)' (Carlo, PytlikZillig, Roesch, & Dienstbier (2009): 277).

[8] 'Observer' and 'victim' are Nichols's terms for the person having the empathy and the person toward whom she has it and I follow his usage in this paper.

In his important discussion of various forms of fellow-feeling, Scheler notes an important moral difference between emotional contagion and fellow-feeling. If we come to be 'infected' by the other's state of mind—say sadness or depression—this can well have the effect of *crowding out* genuine fellow-feeling for the other that is involved in empathy (Scheler (1954): 17). I may become so consumed by or at least stuck in my own sadness caused by the other's sadness that I am unable to muster the other-directed attention and concern required for genuine empathy, or other fellow-feeling for her.

Martin Hoffman, a leading psychologist of empathy, himself provides a different, but still faulty, account of empathy, failing to capture its intentional structure. He (rightly) rejects the common psychologists' (and Nichols's) view of empathy as 'one's feeling matching the other's feeling'. But he puts in its place 'the involvement of psychological processes that make a person have feelings that are more congruent with another person's situation than with his own situation' (2000: 30). He gives the example of feeling empathic anger on seeing someone verbally attacked when the attackee might be feeling sad or disappointed. The anger could be regarded as more appropriate to the victim's situation than is the victim's own sadness, and would not be appropriate to the subject's situation which (let us assume) provides no basis for anger.

But were the observer to feel anger and to believe that this is the emotional response the victim should appropriately have to her own situation, this would not be empathy. The observer's empathy is directed *toward the victim*. It is not a self-standing feeling that is appropriate to the victim's situation. Hoffman's definition implies that the empathizer is in a way confused, having feelings inappropriate to his own situation but appropriate to those of another.[9] This faulty definition does not vitiate Hoffman's many important findings about and speculations regarding empathy, which presuppose that it is indeed essentially directed toward others.

8.4 Mental State or Life Situation as the Target of Fellow-Feeling

So Nichols joins some other philosophers and psychologists in failing to articulate the intentional structure of empathy that helps to bring out its moral significance. Nichols's account of empathy goes wrong in another way as well. This lies in his privileging the victim's state of mind over her actual life situation in what it is that altruistic sentiments are directed towards. This comes out in his discussion of 'mindreading' necessary to activate 'basic altruistic motivation' (2004: 33). I must explain Nichols's background discussion of mindreading before getting to this issue.

[9] Schopenhauer recognizes this very error, in a criticism of the philosopher Cassina: 'His view is that compassion arises from an instantaneous deception of the imagination, since we put ourselves in the position of the sufferer, and have the idea that we are suffering his pains in our person. This is by no means the case; on the contrary, at every moment we remain clearly conscious that he is the sufferer, not we; and it is precisely in his person, not in ours, that we feel the suffering, to our grief and sorrow' (([1841]/1965): 147).

By 'mindreading' Nichols means the observer's recognition or belief that the victim has a particular state of mind. Nichols sets out to defend the view that attribution of distress to the victim—that is, mindreading—is required for altruism. He defends this as a middle position between two others—(a) that no mindreading at all is necessary, and (b) that more sophisticated forms of mindreading (for example, the ability to imagine oneself in the other's situation) are necessary for altruism.

Against the 'no mindreading' view, Nichols cites experiments by Batson, an influential social psychologist researching altruism, in which a subject watches (what she takes to be) shocks administered to another. There are two different scenarios—one in which the subject can easily escape the situation and another in which she is compelled to stay for a more extended period of time.

What Batson finds is that there are some people who will help the distressed person even when they are in the 'easy escape' situation, i.e. when they could relieve their own (i.e. 'personal') distress by escaping from the situation. Nichols reports this result by saying 'so there is an important kind of altruistic motivation that cannot be satisfied by escaping the situation' (2004: 38). Nichols implies that some people act altruistically to help when they could escape, while others act altruistically to help the victim as a way of relieving their own personal distress (and some will escape the situation as a way to relieve their own distress). But those who help the other in order to remove the cause of their own distress are not acting altruistically; they are simply engaged in helping behavior for self-interested reasons. What Batson's experiment shows is that not all helping behavior is motivated by an attempt to relieve the observer's own distress, not that the latter is actually a form of altruistic motivation different from a desire to help the other for her own sake.[10] To put it another way, Batson's experiment suggests that some persons are motivated by empathy to relieve the other's distress even though they could escape the situation without doing so. There is by definition *no* kind of altruistic motivation that can be satisfied by escaping the situation, so an experiment cannot establish that there isn't.[11]

[10] Nichols at one point accepts that altruism refers to a motivational state and not simply an intentional helping of the other. He says, 'Altruistic behavior is one form of prosocial behavior, but prosocial behavior also encompasses selfishly motivated behavior that helps others. For example, intentionally helping a stranger merely to impress onlookers counts as prosocial, though presumably not altruistic' (2004: 31 n. 1). So this definition is out of line with his use of altruism in the quote about Batson's experiment in the text.

[11] In a gloss on Batson's experiment, John Doris and Stephen Stich (whose work is associated with the experimental philosophy movement) say that Batson does not see his result as conclusively showing the existence of altruism.

> [T]here are also a variety of egoistic routes by which empathy might lead to helping behavior. Perhaps the most obvious of these is that empathy might simply be (or cause) an unpleasant experience, and that people are motivated to help because they believe this is the best way to stop the unpleasant experience that is caused by someone else's distress. (Doris & Stich (2006): 23).

If empathy is directed toward the other and carries with it some motivation to help the other, it cannot simply be 'an unpleasant experience' that could motivate the observer to help only as a means of relieving that unpleasantness. Doris and Stich are misled by the same faulty use of Batson's experiments as Nichols is. Perhaps empathy can wane and be replaced by the mere unpleasantness of viewing the other's distress (i.e. personal distress). But empathy *itself* cannot give rise to such personal distress. If the directedness of altruistic concern toward the other is or becomes absent, there is no longer any empathy.

Nichols wants to argue that some mindreading is necessary for altruistic motivation; but he does not need Batson's or any experiment to establish this, nor would they be relevant. 'Minimal mindreading' is required for altruistic motivation by definition. Unless the observer knows that the other is in a state of distress (or need, or possible benefit, etc.), he cannot be motivated by altruism, as altruism requires recognition of the other's state (and a direct motivation to improve it for its own sake). But that recognition precisely *is* mindreading (of at least the minimal form Nichols desires).

Let me return to the issue of Nichols's privileging the victim's occurrent state of mind over her life situation as the object of fellow-feeling. He discusses another challenge to the minimal mindreading account. This is Sober and Wilson's view that sympathy sometimes motivates altruistic behavior; that sympathy is different from empathy; yet that sympathy does not require mindreading. The first point is intuitively plausible, and Nichols does not dispute it. In support of the latter two points, Sober and Wilson construct an example of Walter discovering that, unbeknownst to his friend Wendy, her husband has been unfaithful to her. Walter could sympathize with Wendy even though he is not directing that sympathy to Wendy's state of mind, since she is not aware of her husband's infidelity (Nichols (2004): 38; Sober & Wilson (1998): 234–5). As Sober and Wilson put it, 'You can sympathize with someone just by being moved by their objective situation; you need not consider their subjective state' (Sober & Wilson (1998): 236). So being moved by an objective situation would be a genuine example of altruistic motivation seemingly without any mindreading.

Nichols complains that Sober and Wilson provide no evidence for their 'objective situation' view (Nichols (2004): 39). But there seems to be a significant burden of proof issue that bears on Nichols's privileging of social psychology experiments as the basis for his philosophical claims. For it seems clear that we are quite often moved by the plight of others just from knowing that plight and not from giving thought to, nor perhaps even knowing anything about, their state of mind, nor imagining what their state of mind is.[12] If I receive a fund-raising letter from an international aid organization with a photo of a child suffering from AIDS, I might be moved to give money without thinking about what that child is feeling about her disease. Knowing she has AIDS might well be sufficient to move me to give.

This is true in our more immediate environment as well. I might have sympathy for a colleague, with whom I am not intimate, who I learn has a child with mental health problems. I might express that sympathy if I run into the colleague, or even drop a note to the colleague—yet without giving thought to the colleague's actual state of mind. It is enough that I know that her plight is a bad one for her.[13]

[12] Nichols hypothesizes that the state-of-mind theory can be saved if Walter imagines what Wendy would (be likely to) feel were she to discover the infidelity.

[13] This is not to say that it is *typical* to fail to give any thought to a person's state of mind of whom one only knows her plight. Especially when we are emotionally close to someone, we are likely to do so.

So it certainly seems possible to have sympathy for someone's plight without the 'mindreading'. It would seem to me that the burden is not on Sober and Wilson to produce experimental evidence for this possibility, but for Nichols to show why our ordinary experience of this should not be credited. The argument here is not mere 'armchair' pumping of conceptual intuitions but simply reminding ourselves of a familiar experience (that we are sometimes moved by sympathy for someone's objective situation without giving thought to the occurrent state of mind of our victim). This is a familiar and seemingly indispensable form of philosophical argument; and it seems on as sure a footing as a particular laboratory experiment in psychology. So Sober and Wilson seem correct to claim that we can be moved by sympathy (and perhaps other forms of fellow-feeling as well) by the 'objective situation' of others, not by their occurrent state of mind.

8.5 Young Children and 'Core' Cases of Altruism

Nichols gives a second argument against the importance of sympathy in his project of understanding altruistic motivation that bears on his privileging of the occurrent state of mind. This depends on his use of examples of children's altruism, an important part of his argument. Nichols says he is setting out 'core' cases of altruism (though he says new evidence might lead us to revise what we regard as core cases). He cites these examples: A twelve-month-old child retrieves a cup for a crying friend; a two-year-old accidentally harms a friend, looks concerned, and offers the hurt friend a toy (2004: 34). Referring to these examples (among others) Nichols says, 'even if Sober and Wilson's "sympathy" does exist, they provide no reason to think that it explains anything like the core cases of altruism with which we began' (2004: 39).

There are several problems with this argument. First, Nichols says that 'children only begin exhibiting the characteristic signs of sympathy after the first birthday, and at this age, they probably have some rudimentary mindreading skills' (ibid.). He takes this to show that the capacity for sympathy requires the capacity for mindreading. That is, Nichols treats the view that the general capacity for sympathy requires the general capacity for mindreading as equivalent to or at least as requiring the view (disputed by Sober and Wilson) that on every single occasion on which someone has sympathy for another, that person attributes a particular state of mind to the victim. We can think that a person who was completely unable to recognize what others are feeling is also incapable of sympathy. But this does not entail that on every occasion in which a person has sympathy he attributes a state of mind to the other.

This point leads to a second, which concerns what it means to say that the childhood examples Nichols gives are 'core' examples of altruism that, Nichols says, Sober and Wilson fail to explain. At one point he takes it to mean cases that emerge early in children yet are pervasive among adults, and says that 'comforting and helping others in distress' are core in that sense (2004: 35). But the impression is also given that by 'core' Nichols also means *uncontroversial* cases of altruism. Yet Nichols's examples might be typical of

childhood altruism in this latter sense, while lacking some central features of fully developed adult altruism. Indeed, it seems plausible that as children develop, they become more and more able to differentiate a victim's plight from his current state of mind. So even if adults comfort and help others in distress, their doing so might have important features lacking in childhood cases of such, but also there might well be pervasive features of adult altruism that do not show up in very young children at all. Martin Hoffman, the psychologist of empathy mentioned earlier, puts the point this way:

> At some point in development, owing to the emerging conception of self and others as continuous persons with separate histories and identities, children become aware that others feel joy, anger, sadness, fear, and low esteem not only in the immediate situation but also in the context of a larger pattern of life experience. Consequently, although they continue to feel empathically distressed in response to another's immediate pain or discomfort, they can also respond with empathic distress to what they imagine to be the other's chronically sad or unpleasant life condition. (Hoffman (2000): 80)[14]

For adults, it would seem that the 'core' object of altruistic concern is not the state of mind as such but the victim's plight. The state of mind is of altruistic concern insofar as it is a part of, or a symptom of, her actual or threatened ill-being. Normally the state of mind and the plight form a cohesive whole to which the observer responds with altruistic concern (when appropriate). Typically, our response to others, even when focused on an occurrent state of mind, takes the context of their life situation, as it takes place over time, into account, even if only implicitly. If my friend reports sadness about her boyfriend leaving her, but seems quite cheerful and behaves with a new sense of unburdenedness, I might infer that she is not really sad about her boyfriend leaving her, but only thinks she is; or that she is, but only to a small or superficial degree.[15] But a person's current state of mind can sometimes be out of line with her plight, as Sober and Wilson's earlier example of the wife who does not know her husband is cheating shows. And when that is the case, it is clearly the plight that trumps the state of mind with regard to (appropriate) altruistic concern.[16]

It is plausible to adopt Hoffman's developmental view cited above that part of moral maturity consists in coming to view the other's life situation as the necessary context for a focus on her well-being, and that altruistic concern has this focus. Occurrent distress

[14] Earlier I mentioned Hoffman's potentially misleading use of 'empathic distress' in this passage. (See page 176.)

[15] Eva Johansson (2008) argues that even young children are responding to the shared 'life-world' of other children—to their embodied selves, their socially situated selves, not merely to what one child takes to be another's inner states. (The notion of 'life world' is taken from the phenomenological tradition, especially, in Johansson's case, that of Merleau-Ponty.) The examples below (18) illustrate Johansson's point, showing young children being concerned about another's situation, not simply her occurrent state of mind.

[16] Nevertheless, a state of mind can also be of moral interest independent of the victim's objective situation. If my friend feels a sense of loss about something the loss of which I regard as in her long-term best interest (e.g. what I regard as a destructive relationship for her), it is appropriate for me to feel empathy for her sadness; yet also to have the appropriate fellow-feeling—e.g. relief, gladness—directed toward her overall situation. This is because her sense of loss is itself a genuine hurt she is feeling in the present. But it is still the state of mind as constitutive of well-being that is the appropriate object of fellow-feeling.

such as the children respond to altruistically in Nichols's examples is expanded in later personal development to this larger canvas of the life situation. Indeed, it is worth noting that the examples from children that Nichols has chosen to present are biased in favor of his 'state of mind' account. The article from which he draws the examples contains other examples that have an 'objective situation' rather than a 'state of mind' focus and thus are closer in spirit to the life situation focus. For example, in one a 3½-year-old is playing with his 6-month-old sister on the floor. He sees a safety pin and takes it to his mother, saying that it would hurt his sister if she touched it. In another, a 3-year-old gives a friend her own Donald Duck hat because the friend has (recently, but not at the current moment) lost her own (Boston Celtics) cap. Both these cases involve a young child's response to another's objective situation rather than to a definite current state of mind.[17]

So Nichols has been selective in choosing his experimental examples. A fuller look at those examples does not support his privileging the victim's state of mind over her plight as the 'core' object of fellow-feeling, nor, therefore, his argument for demoting the importance of sympathy in an understanding of altruistic concern (because of seeing it as not 'core').

Following his presentation of the childhood cases to get to adult cases, Nichols says the following: 'The clearest real-life examples of altruistic behavior in adults come from work on helping behavior in social psychology' (2004: 34). To illustrate this, he describes in some detail a laboratory experiment of Clark and Word (1974) (the details are not relevant here) that finds a large percentage of people offering aid when no one is watching them. It is striking that Nichols finds such a laboratory experiment to be the 'clearest real-life example' of altruism. It would seem that the clearest real-life examples are those that come from our lives in our world. We can all produce examples in our own experience of people helping others from what every evidence available to us suggests is altruistic motivation. And some such real-life examples have been extensively studied, for example, rescuers of Jews during the Holocaust (Oliner & Oliner (1988); Monroe (1997)). As real-life examples of altruism they are likely to serve us better as material for reflection than social psychology experiments in laboratory settings. When we engage in that reflection, we are not simply 'sitting in an armchair'. We are drawing on our moral sensibilities, our experience, and our reflective capacities to engage with human phenomena that we recognize simply because we are morally competent human beings living in a world of other human beings.

It is noteworthy that, when Nichols talks about children, he does not cite experimental data but draws on observation of actual children as 'core' cases. It would seem that reflection on adult morality and adult altruism should similarly begin with cases that we can recognize from our world. In the scholarly community of experimental

[17] Blum ([1986]/1994): 187). It is possible that in the second example given the observer infers a kind of distress in the victim; but even if so, the child's life situation is part of the observer's reason for responding the way she does, and Nichols's account of childhood concern does not allow for this.

psychology, the Clark and Word experiments (and Batson's as well) may well constitute important progress, because they challenge a reigning paradigm in that community that all human behavior is ultimately egoistically motivated.[18] But the understanding yielded by these experiments cannot be equated with nor be paradigmatic of an understanding of the character of human altruism in all its complexity, such as philosophical reflection with its distinctive ways of surveying human capabilities and the human condition can provide.

The nub of Nichols's defense of minimal mindreading lies not in refuting the 'no mindreading' view, which I have argued is a non-starter, but in arguing that nothing more robust than minimal mindreading is necessary for altruistic motivation. I think his argument on this point is largely successful and I will present it briefly. At the same time, I think here too, Nichols operates with an impoverished view of fellow-feeling, missing its intentional structure, its focus on the other, its difference from an occurrent affective state, and, more generally, its organic connection to human relationality.

The main form of robust mindreading Nichols considers is 'perspective taking', by which he means imagining oneself as having the mental states of another (2004: 8), or as taking the perspective of another person (2004: 9). These two characterizations are not really equivalent, but I understand Nichols to mean imaginatively projecting oneself into someone else's circumstances as a way of envisioning what she feels. Nichols recognizes that perspective-taking in this sense can be a path to empathy. But his point is that empathy can be reached without it. In particular, it is implausible to think that the children whose empathy he has described have engaged or are capable of engaging in such a cognitively sophisticated process. If we are willing to accept these children as cases of fellow-feeling of some sort, then perspective-taking cannot be a necessary condition for empathy.[19] This argument seems correct to me.

[18] Eisenberg says that in the 1980s 'there was considerable interest [among psychologists] in why people sometimes help others at a cost to themselves and whether truly selfless altruism exists', contrasting this with an earlier period in which egoistic theories of human motivation were more dominant (2000: 672). See Scott & Seglow (2007): ch. 4 for a recent, brief account of the emergence of a recognition of genuine altruism (not merely subtle or disguised forms of pursuit of self-interest) in social psychology. Batson is reported to have done an about-face, originally holding to the egoistic hypothesis (Barasch 2009: 168). An earlier collection, Mansbridge (1990), lamented the dominance of egoistic paradigms in several different disciplines, including psychology, and tracked a then recent emergence of alternative views that allowed for genuine altruism.

[19] Goldie makes the important point that imagining oneself in another's place is, in contrast to empathy, by itself altruistically inert. It involves no necessary concern for the other, but can be engaged in from a variety of motives, including a malicious desire to savor the distress of the other (Goldie (2001): 215). Johansson provides a wonderful example of such imagining going disastrously wrong with young children. A 2½-year-old, Zakarias, is pulling the sweater of a 3-year-old, Sture. The teacher tries to get Zakarias to stop doing so by inviting him to think about how Sture is feeling through remembering when he himself had his sweater pulled by another child, Oskar, who is present. But the effort to remember does not lead Zakarias to empathy for Sture, but only to anger at Oskar (Johansson (2008): 42). None of this is to deny that from the perspective of moral education, encouraging people to see the world from another's point of view is frequently an important way of encouraging fellow-feeling.

8.6 Appreciating the Other's Difference from the Self

Nichols also considers a form of mindreading more minimal than perspective-taking but still more robust than his own view. That is the ability to recognize that others might have different desires than one would have in the same situation. For example, Marisol might be able to ignore a public criticism of her work that would make Tamsin extremely upset were she its target instead of Marisol. Marisol's fellow-feeling should be directed toward what Tamsin actually feels, not what Marisol would feel were she in Tamsin's situation. And some people are quite poor at discerning the difference between what they themselves would feel in a given situation and what others in that situation do feel.

Nichols calls failure to appreciate this sort of distinction between myself and others 'mindreading egocentrism'. He distinguishes it from 'ethical egocentrism', which is a failure to be motivated by the well-being of others for their own sake. And he argues that the former does not entail the latter. That is, someone who keeps using his own needs and desires as a template for understanding others, and thus often misattributes their state of mind, could still be genuinely empathetic and concerned about their welfare. I think this argument is valid.[20] But I also think that Nichols's argument misses something important. He says, 'Of course, if she is an egocentric mindreader, she may not be very effective in helping others, because she will not be sensitive to the variation in desires, feelings, and thoughts that actually exists among those she tries to help' (2004: 51). This makes it seem that mindreading egocentrism is deficient only because of its consequences—it leads to inefficient or ineffective helping. But we often find something defective in the observer's sentiment toward the other itself when it fails adequately to track the actual thoughts and feelings of the other—that is, when it fails to grasp the other's actual situation. The victim may well have reason to feel that the observer is not really caring about her if he takes action that is apparently motivated by what he regards as her good, but is in fact beholden to a mistaken view of her and her good. Recognizing that mindreading egocentrism can be a deficiency of moral response, not only of outcome, is a strength of the 'care ethics' tradition, which emphasizes that caring about someone involves giving attention to her in her individuality and distinctness—seeing her as a distinct person with her own desires, needs, and interests that might well be different from that of the observer.[21] This notion of 'attention' is derived from Iris Murdoch (who in turn got it from Simone Weil), who saw it as central to the moral attitude, fundamental to recognizing the 'reality of other persons', as she put it (Murdoch (1970); Weil (1962)). Mindreading egocentrism

[20] Nichols attributes the 'discrepant desire' view (the view that altruistic motivation requires the ability to distinguish the desires of others from one's own) to me (1994: 183–214, especially 193).

[21] Noddings (1984) was the first in the recent care ethics tradition to emphasize this point, though the care tradition has subsequently been further developed, incorporating and building on her insight. Scheler (1954) was an earlier philosopher who strongly emphasized this point.

is morally problematic not only because of its likely suboptimal consequences but also because it betokens a defective emotional attitude toward the victim.

Indeed, the notion of 'altruism' itself can fail us in this situation precisely because it suggests a purely motivational state severed from its embeddedness in broader sentiments and attitudes toward the other. Nichols is right to distinguish mindreading from ethical egoism, and to say that only the latter contravenes altruism. But mindreading egoism is itself contrary to the larger attitudes and sentiments of care and concern that provide the backdrop of altruistic motivation in an adequate account of moral responsiveness to the plight of others.

That care and concern are attentive to the ways that the victim's needs and desires may differ from the observer's, as well as being aimed at what the observer takes to be the good of the other, is supported by the developmental story Hoffman tells, cited earlier (p. 180). Nichols emphasizes the way that toddlers sometimes offer their own comfort object to a distressed playmate, rather than the playmate's comfort object, as an example of mindreading egocentrism (2004: 51). Yet Nichols's account of his own examples is selective, omitting the child's tracking of the other's well-being that is involved in the example itself, in the source from which Nichols has taken it. He gives the example of Michael, a 15-month-old, struggling with his friend Paul over a toy, causing Paul to cry. Michael seems distressed and brings his own teddy bear and security blanket to Paul—thus, mindreading egocentrism. But what Nichols omits from Hoffman's larger description of this example is that Paul continues to cry and Michael goes into the next room and brings Paul's security blanket to him, implying that he has in some way recognized that his own comfort object will not do the trick for Paul.[22]

In his account of the development of empathy, Hoffman sees 'self/other differentiation' as the key element in that development from infancy to late childhood. For example, he sees 'awareness of self and others as having independent internal states' and 'awareness of self and others as having their own personal histories, identities, and lives beyond the immediate situation' as stages in that development.[23] This is not a point against Nichols's claim that mindreading egocentrism and ethical egocentrism are two different things. It is meant to support the view that we see mature moral responsiveness to others as involving an awareness of the ways that others can differ from the self, an attentiveness to the particularity of the other, and a desire for the other's well-being, all wrapped up together. This complexity of the altruistic emotions cannot be captured by a view of them as simply occurrent affective states with no conceptual/intentional relationship to the human (or sentient) other.

[22] The example, as mentioned, is drawn by Nichols from my essay 'Moral Development and Conceptions of Morality' ([1986]/1994), and I took it from Hoffman (1976).

[23] Hoffman (1976): 64. This is a different point from the other drawn from Hoffman (1976: 80) that moral development involves the ability to differentiate a person's state of mind from her ongoing situation. See p. 180, above.

8.7 Affect and Motivation: Nichols's 'Concern Mechanism'

After the defense of 'minimal mindreading' as necessary (but also sufficient) for the attribution of distress, Nichols says that he still needs to explain how that attribution leads to altruistic motivation: 'I will try to characterize the affective response underlying altruistic motivation' (2004: 52), which he calls the 'Concern Mechanism'. It involves three components—(1) the attribution of distress to the victim; (2) an occurrent affective state produced by the attribution; and (3) such affect motivating the observer to take action to relieve the other's distress. This analytical framework, I will argue, falls afoul of the problems I have noted so far in Nichols's argument: his failure to sustain the morally significant distinction between self-directed and other-directed emotions; his identification of complex altruistic emotions with mere occurrent affective states; and his omission of the intentional character of altruistic emotions. But in addition, this framework distorts both the perceptual and cognitive dimension of altruistic emotions, and the way that actions motivated by such emotions express them and thereby affect the character of the act itself. More generally, the 'Concern Mechanism' framework distorts the relationship between perception, emotion, and action.

By using the word 'concern' for this mechanism in the context of saying that he is investigating altruistic motivation, Nichols implies that the mechanism provides a way of understanding concern for the other, that is, concern for the other's well-being, a genuinely altruistic attitude. We would not ordinarily say that if I help someone out of a desire to relieve my own personal distress (caused by her suffering), I have acted out of 'concern'. Yet when Nichols characterizes the Concern Mechanism, he includes within it 'sympathy' and '2nd-order contagious distress'. Second-order contagious distress is like the contagious distress mentioned earlier (p. 172), in that it is a distress in the observer caused by the distress in the victim. What makes it 'second-order' is that the observer first attributes distress to the victim and it is this attribution that produces the distress in the observer.

This form of contagious distress can motivate the observer to help relieve the victim's distress; but it does so only as a way to relieve the observer's own distress, in the same way first-order personal distress can. So second-order order contagious distress provides a motive to help that is no more altruistic than that constituted by first-order contagious distress. Neither of these is a form of *concern*, because neither is a concern *for the other*. Second-order contagious distress thus does not belong in the same category as sympathy as a form of concern.

A second problem with the Concern Mechanism is its rejection of the view that altruistic motivation can be produced without the intervention of an occurrent affect (step (2) of the Mechanism). I think that I have sometimes been motivated to help from the mere recognition that someone needs help, without the intervention of any

affect—either from the mere thought that the person needs my help or the thought that it is the right thing to do. This seems a fairly familiar phenomenon.[24]

I think Nichols dismisses this possibility because he conflates it with a much stronger one—that a human being would be capable of altruistic motivation without having any capacity for emotional responsiveness to the suffering of others. He attributes this latter view (though not distinguishing it from the former) to Sober and Wilson.[25] I don't think I would be capable of altruistic motivation were I to be a creature without any affective life at all. My ability to be moved by the mere recognition of plight in the absence of occurrent affect still depends on my having emotions, affectively based capacities that *sometimes* involve occurrent affect.

So Nichols is incorrect to think that concern for others, when motivating helping action, always involves an occurrent affect. I can be concerned about Jill and do something to help her without feeling concern (or any other specific feeling toward Jill) at the moment. But more generally, if one appreciates the altruistic emotions as complex states of mind with intentional references to others' well-being and concomitant forms of perception and cognition involved in them, the entire picture of the 'Concern Mechanism'—not simply the claim that occurrent affective states are always required for altruistic motivation—is thrown into question.

In the Mechanism, the observer first engages in an act of attribution of distress to the victim, temporally followed by and causing an affective state. But this is not the usual phenomenology of seeing someone in distress. Typically (if not in every instance, a qualification I will discuss in a moment) to see that someone is in distress already involves fellow-feeling toward her. 'Distress' is not an evaluatively and emotionally neutral attribution to another. Suppose I am at a meeting at work and a co-worker, Hector, is acting oddly. He seems distracted, as if his attention were elsewhere. Then I remember that a second co-worker has told me that Hector has been having some trouble with his teenage daughter that has been causing him some mental anguish. I now come to see his behavior as manifesting distress, which I had not done initially. And in doing so, I feel sympathy or empathy for him (for the example, it does not matter which). Viewing Hector sympathetically and seeing his distress are part of the same mental act for me. Seeing the other's piece of behavior (his staring out the window, or fiddling with his smart phone) as manifesting distress and viewing the other

[24] I am not relying here on the strong form of a Kantian thesis that a purely rational principle is capable of moving us to help others, no matter what we feel at the moment (although I believe that to be true). I think we can be motivated by the mere recognition of another's need or suffering, not mediated by the thought that such an action is prescribed by a principle prescribing right action.

[25] Sober and Wilson apparently make the evolutionist argument that we might have evolved dispositions to help one another in the absence of any emotional capacities (Nichols (2004): 52). I agree with Nichols in rejecting this view, although, again, Nichols equates this view with the view that on every occasion on which we are motivated to help others, we do so by way of an occurrent affective state.

with sympathy normally go together and are part of the same mental phenomenon. Fellow-feeling is an emotion that involves a way of viewing the other; one might say that it reveals aspects of the other's well-being (and ill-being) that are invisible to the unsympathetic person. That is, to have empathy for someone is not simply to have a feeling state; it is to perceive the other in light of the other's state of well-being (although the perception is not necessarily accurate).[26]

P.F. Strawson made a similar point in his influential essay, 'Freedom and Resentment' ([1962]/1974). He noted a set of attitudes toward other persons—for example, resentment, gratitude, forgiveness, anger, love—that arise naturally in interaction and relationships with them, involving responses to the good will or ill will they show us. He calls these 'participant, reactive attitudes'. Strawson contrasts with them a set of attitudes we can take up toward others that psychically remove us from interpersonal involvement with them and view them as something like objects, to be manipulated or controlled (sometimes for their own good). From the participant's point of view, these reactive attitudes are part and parcel of our relationships with others and we cannot envision having such relationships without those attitudes. But we are capable of adopting an objective point of view that inhibits these attitudes and calls forth other ones (bafflement that someone could act that way, concern to make sure the person does not do that same thing again, scientific curiosity about what makes such a person 'tick').

Similarly, I am arguing that our normal mode of engagement with others means that when we experience someone as in distress or suffering, we are seeing her in a compassionate light; we have some sort of fellow-feeling toward her. We can pull back from that kind of engagement with others, enabling us to see their distress from a distance, as it were, for a number of reasons (compassion overload, a desire to avoid the moral pull of the other's distress that we know will accompany our fellow-feeling). But the normal case of interacting with others involves perceiving someone as in distress and having fellow-feeling toward them as bound up with each other, just as experiencing them as being hateful toward us and resenting them for doing so are. In a sense, then, Nichols's notion of 'attribution', which might seem like a perfectly neutral description carrying no philosophical baggage, is misleading in the normal situation of seeing another as in distress. It implies that there is never an intrinsic or internal connection between this viewing the other as in distress and having fellow-feeling for her.

Nichols implicitly replies to this argument by saying that psychopaths and perhaps other evil persons are perfectly capable of seeing someone as in distress without experiencing any fellow-feeling toward that person. So the attribution and the fellow-

[26] On the idea that emotion affects the reality perceived by the observer, see Goldie (2008): 233, and, more generally, Murdoch, Nussbaum (e.g. (2004): 24–31), and others emphasizing the perceptual dimension of emotional experience. I am also influenced by Vetlesen (1994: 154–63); e.g. he says, 'I argue that to "see" someone as suffering is already to have established an emotional bond between myself and the person I "see" suffering' (1994: 159).

feeling must be two different mental events only contingently related. But this does not follow. Just because a psychopath can attribute distress without having any accompanying fellow-feeling does not mean that the attribution and the fellow-feeling are not tied together in a holistic way in the case of non-psychopathic human beings. In this regard it is significant that we regard psychopaths as humanly as well as morally defective beings. Nichols and others in the neo-sentimentalist tradition have pointed out that psychopaths have three deficiencies. They are generally unresponsive to others' distress; they lack motivation to help others (unless they see doing so as in their interest); and they do not make moral judgments. (They make normative judgments, but Nichols makes a compelling case that they are unable to appreciate the force of the 'moral', shown in part in their failure to be able to distinguish between (what we regard as) judgments of moral wrongness and judgments of conventional wrongness.) These deficiencies suggest that psychopaths' responses to distress cannot tell us what the connections between fellow-feeling and seeing distress are among normal persons.

A further consideration calling into question a purely contingent link between seeing someone as suffering and responding to it is the set of expressions we have that suggest different degrees of understanding of another's suffering or distress. We often say someone 'has not taken in', 'has not really understood', or 'understands only superficially' someone else's suffering, if he verbally attributes suffering to the person but fails to respond to it. That is, these expressions imply some sort of failure of the understanding—then a cognitive failure—not only a failure of appropriate emotional response, thereby implying some sort of internal link between cognizing of distress and responding emotionally to it in an appropriate manner.

Nichols may reply to this as he does to a different but related case (2004: 113) that the person understands perfectly well that the other is suffering but just doesn't care. But I think our notion of 'understand' in application to the plight of others does not draw this distinction so sharply; it allows that there are undoubted kinds of understanding that are absent in those who fail to respond with fellow-feeling of some kind (Starkey (2008): 425–54). This is not to deny that there can be a minimal threshold for Nichols's notion of 'attribution' of distress or suffering, according to which one can say the attributor understands but does not care. My point is that this phenomenon can also be expressed in the epistemic language of failing to understand, to really understand, or to have a full understanding, and that Nichols's conception cannot capture this epistemic usage.

8.8 Action Expressive of Emotion

The 'Concern Mechanism' idea runs into other, not unrelated, problems regarding the relationship posited between affect and helping ((2) and (3), on p. 185). That posited relationship is that the affect supplies a motive for the act, but the character of the act itself is not affected by this motive. That implication is also carried by the somewhat antiseptic

terminology of 'prosocial behavior' or 'helping behavior' favored by psychologists. The problem here is that the character of action that expresses particular emotions is (not always but often) very much affected by the particular emotion in question. The action of helping someone out of compassion or concern has quite a different significance from helping someone as a way to relieve a cause of the observer's own distress.[27] It feels like, and is, a different action, both from the observer/agent's and from the victim/recipient's point of view.[28] That is, emotions such as fellow-feeling do not only *motivate* action; they *express themselves* in action.

The image of a 'mechanism' (the 'Concern Mechanism') tends to mask this important point about the relation of emotion and action by implying a too-sharp divide between the motive/emotion as a purely inner psychological event and the external behavior it is taken to prompt. This problem with the 'mechanism' imagery is comparable to the previously discussed one of making a too-sharp divide between seeing distress and emotionally responding to it. In both cases the oversimplified image of the emotion of fellow-feeling as an occurrent, short-lived, affective event is at fault. Once we recognize the perceptual, cognitive, intentional, and expressive dimensions of altruistic emotions, we will be less attracted to the metaphor of a 'mechanism'.

8.9 Nichols's 'Sentimental Rules'

Recall that Nichols's discussion of responses to distress is in service of a theory of moral judgment. 'Core moral judgment depends on two mechanisms, then, a normative theory prohibiting harming others, and some affective mechanism that is activated by suffering in others' (2004: 18). In contrast to Hume, for whom sympathy (the emotion on which moral judgment rested) is a moral emotion, for Nichols the emotions that are implicated in moral judgment are not moral emotions. 'None of the affective mechanisms explored in this chapter counts as a moral sense' (2004: 62), by which he means a sense that tracks vice and virtue in persons and acts. Nichols takes the argument of his chapter 2 to show that non-psychopathic persons have strong emotional reactions to others' distress, and this powerful emotional reaction attaches to the norms prohibiting causing distress, providing them with a distinctive status (in contrast, for example, to mere conventional norms) that constitutes what we understand as 'moral'. They 'supply the sentiment to moral judgment' (2004: 63).

[27] This is not always so, because in some situations we may care only that a certain result ensues from an action, not what sentiment or emotion the action expresses. For example, if I am being held up at gunpoint and a stranger rescues me by disarming the gunman, I may not care whether the stranger is motivated by compassion or to win a bet with his friend that he could accomplish this rescue effort. (Even in this case, however, I think the motive would make *some* difference to me. I would appreciate the rescuer in a different way if I thought his action expressed compassion or helpfulness than if I thought he was motivated by gain.)

[28] This point was made by Bernard Williams in (Williams [1966]/1973), in saying that the recipient sometimes desires the 'human gesture' from the agent; but the contrast there was with action done out of duty. Blum elaborated Williams's point in Blum (1980): ch. 7. Goldie makes a similar point in Goldie (2001): 41.

This view seems to me to have contributed to Nichols's failure to sustain an appreciation of the moral difference between altruistic emotions and other, egoistic, responses to distress and suffering. He is indeed correct to say that altruistic emotions do not constitute a moral sense as he understands it. Compassion or empathy for another is a response to (or, as I have argued, reveals) suffering; but it does not, or at least not necessarily, involve a moral judgment about the rightness or wrongness of an action. (Nichols rightly cites in support of this view that children younger than two are capable of altruistic behavior but not of moral judgment.) Nevertheless, there is a vast moral difference between compassion/empathy and the other egoistic responses to suffering that Nichols discusses—contagious distress, personal distress, second-order contagious distress, and enduring representation.[29] One is a morally worthy response, the other a morally neutral one, everything else being equal.

Nichols sometimes recognizes this morally significant difference, as we have seen, but only fitfully. The Schopenhauerian tradition would have helped him out, with its insistent focus on the morality of motives, not of judgments. But the focus on the strength of the emotional response, and not its moral character, contributes to the problem. (But that focus is what is distinctive of Nichols's view among neo-sentimentalists.)[30]

The failure to see the distinction between egoistic and altruistic responses to distress vitiates Nichols's claim to have provided an explanation of *moral* judgment. That a norm prohibits a type of action that tends to cause strong feelings in observers does not seem to account for its being regarded as a *moral* norm, which would seem to depend not only on the strength of the feelings but also on their character. Suppose all altruistic emotions, emotions tied to concern about the well-being of others, disappeared in human beings, but emotions (such as contagious and personal distress) remained. That is, individuals' distress caused other people distress but no one had any concern for any one else's distress for its own sake. In such a situation, how would norms arise prohibiting the causing of distress? It would seem that the same result—discouraging the causing of personal distress—could be achieved by norms that prevented or discouraged people in distress from contact with others not in distress.

Beyond that, even if the sorts of norms Nichols is envisioning arose, it is not clear that persons so constituted would be capable of regarding the norms in question as moral ones. Nichols recognizes that it is possible to recognize the validity of a norm that others think of as moral without recognizing the moral character of that norm. He thinks that psychopaths recognize that it is wrong to harm others, but think that the reason for its wrongness is social convention rather than morality ((2004: 13 n. 4) and

[29] 'Enduring representation' is when the observer withdraws from the scene that contains the cause of his distress but doing so does not rid him of a mental representation of that distress, which thus continues to distress him.

[30] 'Norms prohibiting actions that are likely to elicit strong negative affect will be treated as distinctively wrong' (Nichols (2004): 63f.). In a later article, Nichols summarizes the position: '[O]ur natural emotional reactions to harming others confer a special force on the rule against harming others, and this connection with emotion likely played a critical role in the cultural success of the rule' (Gill & Nichols (2008)).

elsewhere).[31] In that spirit, I would question whether someone who was completely incapable of altruistic emotion, that is, whose only emotional reactions to the distress of others are egoistic in character, would be able to see the causing of distress to others as morally wrong. If people's emotions were confined to self-oriented ones, it is not clear that these could underpin or sustain moral judgment. So Nichols's sentimental rules view fails to give an adequate account of moral judgment.[32]

Elements of Nichols's argument point more toward Schopenhauer's than to the Humean one that Nichols favors. That is, compassion is (for Schopenhauer) the fundamental moral motive not because it issues from a moral judgment, but because it is directed toward the weal and woe of others for their own sake. And the evidence that Nichols cites from children provides part of a plausible developmental story as to how human beings develop the capacity for compassion. Children are able to respond in a caring way toward others before they make moral judgments.[33] And adults, while unlike very young children in being able to make moral judgments, are capable of acting morally without invoking, even implicitly, moral norms in doing so. That is, to follow Schopenhauer, they can act with compassion, thereby doing the moral thing, without invoking moral norms. And this is how emotion comes into morality; humans will not be capable of that moral responsiveness without emotional capacities. So emotion is necessary for morality, but not in the way that Nichols postulates.

Nichols and so many other contemporary moral theorists assume that the ultimate object of a theory of morality is an explanation of moral *judgment*. But why not see as an equally worthy object a theory of moral responsiveness—something that includes moral emotions and moral motives, and actions animated by them? With such a perspective, one would be much less likely to be misled into short-changing the moral distinction between personal or contagious distress and genuine concern for others.

More generally, I have argued that Nichols operates with an impoverished view of altruistic emotions (empathy, compassion, concern) as mere short-term, affective states of mind. He fails to recognize their essentially intentional structure, that they are directed toward others in light of their well-being; their essential connection to perception and cognition; and their expressive dimension that is central to their moral significance. Although I have not attempted to demonstrate this, I suspect that

[31] From Nichols's discussion of the basis on which psychopaths say that a type of action is wrong, I am not sure whether the psychopaths he discusses actually regarded the action as wrong (though for reasons of convention) or that they merely recognized that others thought the actions were to be discouraged and expressed this by saying they were 'wrong'. That is, I am not sure that psychopaths understood, or applied, a concept of 'wrong' for any actions that were not in their own perceived self-interest. I am not sure that psychopaths did not take 'wrong' to mean 'socially disapproved of'.

[32] This argument draws on Jones (2006).

[33] Nichols sometimes implies that the child's capacity for emotional contagion is a developmentally necessary stage for empathy and concern; but the psychologist Zahn-Waxler argues that children are able to exhibit concern for others apart from emotional contagion, and do not need to have gone through a stage of contagion (Zahn-Waxler et al. (1992)).

Nichols's love affair with experimental social psychology has led him to retreat from the rich tradition of philosophical moral psychology.[34]

References

Barasch, Marc Ian (2009) *The Compassionate Life: Walking the Path of Kindness*. San Francisco: Berrett-Koehler.

Blackburn, Simon (1998) *Ruling Passions*. Oxford: Oxford University Press.

Blum, Lawrence (1980) *Friendship, Altruism, and Morality*. London: Routledge & Kegan Paul.

—— ([1986]/1994) 'Moral Development and Conceptions of Morality'. In Blum (1994): pp. 30–62.

—— (1994) *Moral Perception and Particularity*. Cambridge: Cambridge University Press.

Carlo, G., PytlikZillig, L., Roesch, S., & Dienstbier, R. (2009) 'The Elusive Altruist: The Psychological Study of the Altruistic Personality'. In: *Personality, Identity, and Character*, Darcia Narvaez and Daniel Lapsley (eds), pp. 271–94. Cambridge: Cambridge University Press.

Clark, R. & Word, L. (1974) 'Where is the apathetic bystander? Situational characteristics of the emergency', *Journal of Personality and Social Psychology* 29: 279–87.

D'Arms, Justin & Jacobson, Daniel (1994) 'Expressivism, Morality, and the Emotions', *Ethics* 104: 739–63.

Doris, John & Stich, Stephen (2006) 'Moral Psychology: Empirical Approaches', *Stanford Encyclopedia of Philosophy*, Edward Zalta (ed.) http://plato.stanford.edu/entries/moral-psych-emp

Eisenberg, Nancy (2000) 'Emotion, Regulation, and Moral Development', *Annual Review of Psychology* 51: 665–97.

Gibbard, Allan (1990) *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.

Gill, Michael & Nichols, Shaun (2008) 'Sentimental Pluralism: Moral Psychology and Philosophical Ethics', *Philosophical Issues* 18(1): 143–63.

Goldie, Peter (2001) *The Emotions*. Oxford: Oxford University Press.

—— (2008) 'Getting Feelings into Emotional Experience in the Right Way', *Emotion Review* 1 (3): 232–9.

Hoffman, Martin (1976) 'Empathy, Role-taking, Guilt, and Development of Altruistic Motives'. In: *Moral Development and Behavior*, T. Lickona (ed.), pp. 124–43. New York: Holt, Rinehart, and Winston.

—— (2000) *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge: Cambridge University Press.

Johansson, Eva (2008) 'Empathy or Intersubjectivity? Understanding the Origins of Morality in Young Children', *Studies in the Philosophy of Education* 27: 33–47.

Jones, Karen (2006) 'Metaethics and Emotions Research: A Response to Prinz', *Philosophical Explorations* 9(1): 45–53.

Mansbridge, Jane (1990) (ed.) *Beyond Self-Interest*. Chicago: Chicago University Press.

Monroe, Kristen Renwick (1997) *The Heart of Altruism*. Princeton, NJ: Princeton University Press.

Murdoch, Iris (1970) *The Sovereignty of Good*. London: Routledge.

[34] I wish to thank Paul Harris, Steven Levine, Kelso Cratsley, Jennifer Radden, and Peter Goldie for comments on earlier drafts or presentations of this article.

Nichols, Shaun (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford: Oxford University Press.

Noddings, Nel (1984) *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley: University of California Press.

Nussbaum, Martha C. (2004) *Hiding From Humanity*. Princeton, NJ: Princeton University Press.

Oliner, Samuel & Oliner, Pearl (1988) *The Altruistic Personality*. New York: The Free Press.

Scheler, Max (1954) *The Nature of Sympathy*. New Haven: Yale University Press.

Schopenhauer, Arthur ([1841]/1965) *On the Basis of Morality*. New York: Bobbs-Merrill.

Scott, Niall & Seglow, Jonathan (2007) *Altruism*. Maidenhead, England: Open University Press.

Snow, Nancy (1991) 'Compassion', *American Philosophical Quarterly* 28: 195–205.

Sober, Elliott & Wilson, David Sloan (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

Starkey, Charles (2008) 'Emotion and Full Understanding', *Ethical Theory and Moral Practice* 11 (4): 425–54.

Strawson, Peter F. ([1962]/1974) 'Freedom and Resentment', *Freedom and Resentment and Other Essays*. London: Methuen.

Vetlesen, Arne Johan (1994) *Perception, Empathy, and Judgment: An Inquiry into the Preconditions of Moral Performance*. Pennsylvania: Pennsylvania State University Press.

Weil, Simone (1962) 'Human Personality'. In her *Selected Essays 1934–1943*, pp. 9–34. New York: Oxford University Press.

Williams, Bernard ([1966]/1973) 'Morality and the Emotions'. In his *Problems of the Self: Philosophical Papers 1956–1972*, pp. 207–29. Cambridge: Cambridge University Press.

Zahn-Waxler, Carolyn, Radke-Yarrow, Marian, Wagner, Elizabeth, & Chapman, Michael (1992) 'Development of Concern for Others', *Developmental Psychology*, 28: 126–36.