EMPATHY AND EVALUATIVE INQUIRY

JUSTIN D'ARMS*

The emotions are back in fashion in academic and intellectual circles. The declining influence of behaviorism in the social sciences and of positivism in legal and philosophical scholarship has rendered these disciplines more hospitable to explorations of emotional phenomena within various theoretical frameworks. Emotions are celebrated not only as necessary constituents of a psychologically healthy human life, but also as expressions of, rather than impediments to, our nature as rational animals. Against the thought that emotions systematically distort our sense of what matters and why, it is now often said that emotional reactions reflect or embody our most important evaluative commitments. Indeed emotions are held by some contemporary writers to reveal, and perhaps even to constitute, distinctive forms of value. Humanists, legal scholars, and social scientists are also increasingly interested in the emotions as motivators and regulators of moral behavior, as well as of various other kinds of social interaction that make communal life possible.

* Professor of Philosophy, Ohio State University. This Article has benefitted greatly from discussion with the participants in a seminar at Ohio State University, and from written comments by John Doris, Daniel Jacobson, David Merli, Kathleen Schmidt, and, especially, Heidi Li Feldman. I am also grateful to Elizabeth Anderson and David Hills for comments on an earlier Article treating some of these ideas.


3. This interest has been especially prominent among (though by no means confined to) feminists and other authors concerned with exploring the idea that emotions have been under-emphasized in previous thought about society and morality because they have been conceived at once as feminine and as irrational. See, e.g., ANNETTE BAIER, MORAL PREJUDICES: ESSAYS ON ETHICS (1994); CAROL GILLIGAN, IN A DIFFERENT VOICE 1-3 (1982); ALISON JAGGAR, FEMINIST POLITICS AND HUMAN NATURE 114-15 (1983); NEL NODDINGS, CARING: A
This renewed interest in emotion has now expanded to include economists and evolutionary psychologists exploring the adaptive role of emotions in social conflict and cooperation.4

Of course, the historical distinction between reason and passion and the allied idea that the conditions for good legal and ethical judgment are cool and dispassionate, is not wholly without merit. It should be granted that emotions can distort good judgment. Sometimes things are not as they seem to us when we are in the grip of an emotional reaction. Not everything we are ashamed of is really shameful, not everything that angers us is really objectionable.

Let us call such cases instances of emotional “error.” The possibility of emotional error shows that it would be a mistake to trust one's emotional reactions in every case. But it is important to recognize that the mere possibility of emotional error is an insufficient basis for returning to an older conception of judgment and evaluation that aspires to immunize these processes from emotional influence. After all, sensory error is possible as well. But to grant that vision and hearing can produce illusions and mistaken beliefs is not to grant that these sensory modalities have no role to play in rational judgment. Furthermore, evaluative judgments in particular seem to be so intimately connected to our emotional reactions to the world that it is far from obvious how these forms of judgment could proceed without emotional influence. The possibility of emotional error should be treated as a reason for caution and fallibilism about our emotional reactions, rather than as a reason to attempt to purge them completely.

This Article focuses on the phenomenon of empathy in light of the increasingly popular ideas sketched above. Any attempt to understand the role of emotional responses in human life should include a study of the ways in which different people's emotions are brought into harmony or conflict with one another. Propensities for and susceptibilities to mutual influence are crucial to successful social interactions, as well as to our collective and individual reflection on what sorts of lives are worth living and why. There are a great many


possible causes of such emotional convergence and divergence. The most interesting and systematic of these derive from empathy. But despite their importance for social life, empathic influences on our emotional and evaluative reactions are often regarded as irrational. Broadly stated, the goal of this Article is to reconcile the idea that emotions play a salutary role in evaluative thought with the phenomenon of empathy.

It has been suggested by a number of writers over the years that empathy is important as a precursor to and motivator of moral behavior. By producing "emotional understanding" of the plight of others, rather than mere "intellectual understanding," the suggestion goes, empathy induces us to care about that plight, rather than ignore it. There is, I think, something importantly correct about this idea of empathy as a source of altruistic or "prosocial" motivation, though articulating it is more delicate than has sometimes been thought. But the useful project of exploring empathy's role in the motivation of human action is already well under way, and I do not propose to contribute further to it here. I am interested in articulating a rather different role for empathy; one that I think is equally interesting, if somewhat more recondite.

This Article focuses on the epistemological function of empathy, rather than its capacity to motivate—which is to say that I will investigate empathy's role as a means of acquiring knowledge or justified beliefs. One evident way in which empathy could be a source of knowledge is by helping us find out how another person feels. By empathizing with someone's feelings, one might come to know something about how she feels that one didn't know before. The claim that empathy can produce such knowledge has been a centerpiece of one central research tradition in the psychological and philosophical literature on empathy: the "simulationist" or "perspective-taking" approach. We will consider this approach in some detail in what follows, along with an alternative


6. Whether or not empathy can confer such knowledge depends in part on how reliable its mechanisms are in providing accurate information about the states of others. We will attend further to this question in due course.
model of empathy: the “contagion” model. I will not attempt to
defend the claim that either of these processes is sufficiently reliable
to produce knowledge. Rather, I want to suggest some further
consequences that might follow if they were tolerably reliable. I do
hope, by contrasting these mechanisms of empathic transmission, to
get clearer about how they might produce knowledge of other minds.
Ultimately, though, this discussion serves a rather different purpose.
I want to argue that empathy is not merely a device for acquiring
knowledge about the contents of other minds, it is a device for
learning about what matters and why. Some empathy plays a
distinctive and previously unrecognized role in the epistemology of
value by providing us with vicarious access to evaluative perspectives
different from our own.

I begin, in Section one, by briefly explaining some of the
important connections between emotions and values suggested above.
In Section two I address empathy, distinguishing it from sympathy,
and offering a general characterization of empathic mechanisms.
More detail is provided in Section three, where I develop accounts of
some specific empathic mechanisms, reviewing some recent research
on the subject. Finally, in Section four I ask how we should think
about empathic transmission, in light of the forgoing discussion. I
argue for what I take to be a surprising conclusion: the comparatively
unreflective and involuntary mechanisms of contagion serve a more
important role in evaluative thought than do the apparently more
rational procedures of simulation.

I. EMOTION AND VALUE

There is plainly some deep intimacy between valuing things (or
thinking that they are valuable) and feeling certain sorts of emotional
responses to them. We commonly feel guilty when we have done
something wrong, ashamed of things we think shameful (or
contemptuous of them, in others), amused at what’s funny, indignant
at injustice, and so on. When things lose their power to elicit any
emotional responses from us, it is natural to wonder whether we still
value them as we once did. But it is not obvious whether or how
these observations can be parlayed into general connections between
emotions and values (or disvalues—hereafter I intend talk of “value”
and “values” to include things with negative as well as positive value).
I want to focus attention on two connections between emotion and
value: one epistemological and the other conceptual. I do not aspire
to establish the existence of these connections conclusively here, but I hope to say enough to make them seem plausible as starting assumptions.

The first assumption is that we are justified in according some defeasible presumption of warrant to our emotional reactions. To understand this suggestion, begin by noting that emotional reactions are typically experienced as sensitivities to evaluative features of the world. Unlike mere feelings (such as a tickle or an itch), emotions are ways of taking things to matter. Because they purport to be responses to some kind of value, emotions are amenable to forms of rational assessment that mere feelings are not. In being amused, for instance, we take ourselves to be reacting to something that is *funny*. The assumption of warrant claims that, absent countervailing considerations, we are entitled to count the fact that we are amused by something as a reason to think it is funny. Similarly, if you felt offended by someone's behavior, that is a reason for thinking it is genuinely offensive. To grant our emotions this defeasible authority is not trivial. One might suppose that just what it is for the joke to be funny is for it to amuse you, and what it is for the behavior to be offensive is for it to offend you. In that case it *would* be trivial that we should grant authority to our reactions—and not merely defeasible authority.

But a moment's reflection on cases of emotional error shows that this conclusion would be hasty. After all, we sometimes admit that we were mistaken or unjustified in being amused or offended by something. I might have been amused because I was feeling giddy, not because the joke was really funny. You might feel offended because you are feeling defensive for independent reasons, and not because the behavior that now bothers you is really offensive. So we shouldn't take our emotional reactions as "proof" of the evaluative features to which we take them to respond. My assumption is more modest: it makes sense to accord *some* presumptive authority on questions of value to our actual emotional reactions. These grounds can be outweighed, or even utterly defeated, by competing considerations that convince us to discount or disregard the "testimony" of any given emotional reaction. Giddiness or defensiveness, for instance, would be an obvious example of such competing considerations. But absent such further considerations, we naturally and (I submit) reasonably take our reactions as reasons to judge that something is funny, shameful, sad, and so on.

The second assumption is that, for a certain subset of our familiar
evaluative categories, to think that something is valuable is to think that certain emotional reactions to it are appropriate, in a special sense of *appropriate*. Call this conception of the relationship between emotion and certain values "sentimentalism." Most sentimentalists believe that all evaluative concepts can be explained as assessments of the appropriateness of emotional responses. I think that view is overly optimistic. But there is an important range of values that seem to carry an especially intimate tie to human emotional responses, and for which sentimentalism seems the only plausible account. Consider, for instance, the following concepts: shameful, fearsome, enviable, disgusting, funny, and pitiful. These are clearly evaluative concepts. Each seems to be affiliated with a specific and familiar emotional response (in many cases the term for the concept is a cognate of the name for such a response). For each of these concepts, I submit, to think that something has the feature in question is to think that the paired emotional response is appropriate. So to think something funny is to think amusement appropriate, to think something shameful is to think shame appropriate, and so on.

On the sentimentalist view, then, actually feeling the relevant response is neither necessary nor sufficient for making the judgment.

---

7. Contemporary sentimentalists include ELIZABETH ANDERSON, *VALUE IN ETHICS AND ECONOMICS* 1-5 (1993); SIMON BLACKBURN, *ESSAYS IN QUASI-REALISM* 15-17 (1993); SIMON BLACKBURN, *RULING PASSIONS* 1-3 (1998); ALLAN GIBBARD, *WISE CHOICES, APT FEELINGS* 36-38 (1990); John McDowell, *VALUES AND SECONDARY PROPERTIES*, in *MORALITY AND OBJECTIVITY* 110, 111 (Ted Honderich ed., 1985); and DAVID WIGGINS, *A SENSIBLE SUBJECTIVISM*, in *NEEDS, VALUES, TRUTH: ESSAYS IN THE PHILOSOPHY OF VALUE* 185, 185-87 (1987). There are important differences between these versions of sentimentalism which I will largely ignore in the discussion that follows. The view I favor is called "rational sentimentalism," which aims to explicate the relevant notion of appropriateness in terms of a "fit" between the emotion and the circumstances. Daniel Jacobson and I have been developing this theory in recent collaborative work. Much of this section abbreviates arguments made in this collaborative work. Philosophers take note that this view is not to be confused with dispositionalism. Dispositionalist accounts of the concepts on which I focus here are untenable, as we argue in Justin D'Arms & Daniel Jacobson, *SENTIMENT AND VALUE, IN ETHICS* (forthcoming).

8. If this be doubted, consider that to apply any of these concepts to something is to consider it good or bad in some respect. More specifically, they are apparently "thick" evaluative concepts, in that they serve not only to praise or condemn, but also to describe. Cf. BERNARD WILLIAMS, *ETHICS AND THE LIMITS OF PHILOSOPHY* 129-30 (1985). Also, each of them seems to supply one with reasons: reasons to feel, at least, and perhaps to act as well. Thus, if one thinks a trait shameful, for instance, one thinks its bearer has a reason to be ashamed of it, and one may well think he has a reason to eliminate or conceal it, too.

9. What's the alternative? Can it seriously be supposed that there is some property "the disgusting" or "the funny" that one can understand and attribute to things without thereby involving oneself in judgments about the appropriateness of amusement or disgust? What property might this be, and by what means would we hope to discover its contours?
One can feel ashamed of a trait while thinking the trait isn’t really shameful. This would be to think that the shame one is feeling is inappropriate. (Think of a gay teenager who is convinced that he shouldn’t be ashamed of his sexual orientation, but can’t help feeling that way, at least in certain company.) One can also judge something disgusting without always mustering the disgust that one thereby deems appropriate. (Think of a detective who has been searching the dumpster for clues long enough that he has become inured to the stench—he might still think that this is a disgusting task.) But the sentimentalist claims it would be a mistake to conclude from the possibility of such differences between judgment and response that the relevant values are independent of our emotional susceptibilities. Rather, the point of these evaluative concepts is to focus our thought about what things do and do not merit feelings of this kind in order to regulate these feelings in accordance with norms that make sense of them. So we have these concepts because we have these responses and a capacity to reflect on them. Nothing would count as a genuine application of such a concept unless it involved an endorsement (of the relevant sort) of having the paired emotional response. Call the evaluative concepts that possess this especially intimate sort of connection to emotional responses “response-dependent concepts.” The second assumption, then, amounts to the claim that some evaluative concepts are response-dependent in this sense.

The two assumptions are independent claims; but they interact. The presumption of warrant for our emotional reactions is most plausible with respect to response-dependent concepts. After all, thought about what things have value must begin somewhere, and surely when the values in question are response-dependent, it begins with our actual responses. But the first assumption applies more widely as well: whenever an emotion presents itself as a sensitivity to some aspect of the circumstances, it makes sense to accord it an initial presumption of warrant. “Taking offense” (or “feeling offended”), for instance, is a familiar emotional reaction that often precedes or follows upon a judgment that one has been mistreated in some way. It may be to some extent culturally idiosyncratic—its nature determined in part by particular conceptions of what kind of

10. This way of putting the point, and indeed much of the surrounding discussion, is indebted to Gibbard, supra note 7, at 55-57.

11. Both Gibbard, supra note 7, and Wiggins, supra note 7, make use of speculations about the origins of our evaluative concepts in the sentiments, though in quite different ways.
treatment is due to people and why. It may be that, for this reason, “mistreatment” is not helpfully regarded as a response-dependent concept in the sense above—the response itself already presupposes the concept it would be invoked to explain.\textsuperscript{12} Still, one way in which we come to regard someone as having mistreated us is through feeling offended, and then finding, upon reflection, that we think these feelings are appropriate. Much the same could be said of many evaluative judgments: we are sometimes led to make them by having an involuntary emotional response (of a more or less basic sort) to something, and finding this response reasonable upon reflection. Deciding what is good or bad, praiseworthy or objectionable in various ways is often a matter of having various emotional reactions and reflecting on their appropriateness.

“Evaluative inquiry” is inquiry into what has value.\textsuperscript{13} Much of this inquiry, I’ve been suggesting, naturally arises out of emotional reactions and our reflection on their appropriateness. In feeling ashamed of something, we “feel as though” it were shameful.\textsuperscript{14} In asking whether it really is shameful, we take a critical distance from our immediate reactions and question the evaluative convictions they urge upon us. Such questions, though, should not typically be understood as asking about the justification of evaluative convictions as against nihilism or skepticism about values. Evaluative inquiry typically proceeds against the background assumption that some things are and others aren’t shameful (funny, lewd, wrong, beautiful, important, etc.). The task of this inquiry is to decide which things

\textsuperscript{12} Sentimentalists disagree about whether emotions must be prior to the response-dependent concepts they are invoked to explain. I cannot address that question here.

\textsuperscript{13} All this talk of values in the world, of emotions being appropriate or making mistakes about these values, and, in general, of reflection on questions of value can sound naïvely realistic about claims whose objective status is, to say the least, problematic. But I can’t enter here into debates about whether there really are truths or facts about value; and I do not think any such discussion is required. The fact is that we all routinely talk about what is and isn’t funny, shameful, enviable, disgusting, etc. We sometimes allow that things are funny despite our failing to muster the relevant emotional response, or that they aren’t really funny even though we find ourselves amused. Thinking and talking in such terms are so central to human experience and social life that it is difficult to see how we could proceed without them. So, any philosophical conclusions we might reach about the status of such thought would have to leave us room to continue engaging in it as emotional participants in an evaluative life, even if we rejected it as theorists. In recent years, even philosophers who doubt whether there are truths about value, or who suspect that all evaluative judgments are in error, have recognized a need to leave a role for judgments about value in our discursive practices. Cf. BLACKBURN, supra note 7, at 15-17; GIBBARD, supra note 7, at 81-83; JOHN L. MACKIE, ETHICS, INVENTING RIGHT AND WRONG 15-17 (1977).

\textsuperscript{14} See GREENSPAN, supra note 1, at 148.
have which features. The kind of justification we aim for in such evaluative inquiry is best understood as relative to a contrast class that includes competing evaluative convictions, but does not include (and hence does not purport to answer) thorough-going skepticism or nihilism.\(^1\)

This refusal to regard evaluative inquiry as inevitably attempting (unsuccessfully) to refute skepticism about value should not lead us to evaluative complacency, however. As agents with choices to make, we still want and need to reflect upon which things are valuable in various ways. An important part of that reflection will involve having and assessing the appropriateness of various emotional reactions. These emotional reactions can make better or worse sense in light of our other evaluative commitments, our views about our reasons for acting in various ways, and our wider patterns of emotional reaction. In some cases, we deem our reactions inappropriate because of their failure to fit with these wider evaluative convictions. In other cases, though, emotional experience can initiate a change in those other convictions. When we find ourselves persistently feeling emotions that do not line up with our considered views, we are inclined to revisit and sometimes revise those views.\(^6\) This role for emotional experience is important not only negatively, as a corrective to evaluative convictions we come to see as mistaken, but also positively, as a source of input that can expand our system of convictions to make a place for new sources of value and interest. For example, many people find that responding emotionally to nonhuman animals leads them to what they see as an improved appreciation of ethical considerations concerning the treatment of animals. By reacting emotionally, they are led to see new sources or kinds of value in animal lives.\(^7\)

\(^{15}\) This line of approach to the epistemology of value is developed by Walter Sinnott-Armstrong, *Moral Skepticism and Justification*, in *MORAL KNOWLEDGE* 3, 14-17, 20-25 (Walter Sinnott-Armstrong & Mark Timmons eds., 1996), to which I am indebted.

\(^{16}\) Of course the mere fact of persistence does not establish that the reactions are to be accommodated, rather than discounted. Sometimes even persistent emotional reactions are best regarded as errors to be resisted. But in a reasonably healthy psychology, persistent emotional recalcitrance is often grounds for evaluative reform. Note that sometimes the best conclusion is that the reaction is appropriate, in the sense that the circumstance really is funny, shameful, etc., but that other evaluative features of the situation are more important for determining how to respond, or (especially) how to act.

\(^{17}\) Again, the point is not that all such changes are improvements in ethical sensibility. Perhaps some are and some are not. The point is rather to recognize the ways in which emotional reactions offer new inputs into evaluative reflection.
So far my emphasis has been largely on how individuals attempt to integrate emotional reactions with more reflective forms of evaluative thought. But my interest in empathy arises in part from the conviction that social influences on our reactions play an important role in determining how we feel. If, as I've been saying, how we feel plays an important role in evaluative inquiry, then social influences on our reactions turn out to have wide evaluative significance. Inquiry and reflection are often discussed as though they were activities an individual undertakes alone: the traditional paragons of intellectual virtue are Descartes and Holmes, uncovering the truth through pure reason. But while the model of rigorous independent thinking has attractions, of course it is largely mythological—especially with respect to evaluative subject matters. On questions of value collective deliberation is the norm: we test our own emotional reactions by comparing them with other people’s. We argue and remonstrate, concede and retrench, hoping to influence one another, but ultimately aiming to achieve some kind of convergence in evaluative judgment. Sometimes logic and broader forms of coherence win the day, but sometimes they give ground to widely shared, emotionally grounded conviction. There are a few evaluative iconoclasts, of course. But for almost all of us, finding a sustainable set of evaluative convictions requires integrating one’s commitments into a community of sufficiently like-minded thinkers. This is hardly surprising, since much of the point of evaluation is to find standards that will govern us in social lives we must live together. But it is all too common to regard the reasons for giving some ground to others as merely pragmatic, and to suppose that the search for the truth should proceed by attempting to abstract away from such influences. The view I am urging is rather different. Though we must and do accord a special authority to our own sensibilities, nonetheless, to the extent that we take seriously the idea that emotional reactions can do better and worse at responding to the values that things have, we must acknowledge the possibility that our own reactions might be improved, rather than always degraded, through the influence of others.\(^\text{18}\)

As beings who are capable of valuing and (sometimes) acting in

---

18. Of course, it is sometimes reasonable to reject such influences, for instance when we have an explanation of the others’ different reactions which give us grounds for thinking that they, rather than we, are in error. But absent any such story, we have no grounds for according fundamental authority only to our own reactions. *Cf.* Gibbard, *supra* note 7, at 171-203.
accordance with what we value, we are naturally interested in deciding what is worth valuing (i.e., what is valuable). Although I have been saying that we do and should accord some prima facie warrant to our own reactions and those of other people, we must also countenance the possibility that these reactions are mistaken, perhaps even systematically so. But how then can we decide which of our reactions to trust, except by appeal to other evaluative convictions which may themselves be misguided? To what extent should we trust our actual emotional tendencies as guides to what is valuable and disvaluable? I do not believe that there is any general answer to these questions. But I will broach (without attempting to defend them here) three principles which seem sufficiently robust to deserve our allegiance, unless and until we find better. First, in general, the more nonevaluative knowledge a person has about a circumstance, the more reason there is to trust her emotional reactions as guides to evaluation. Second, in general, our emotional reactions are more trustworthy to the extent that they are informed by a wider range of life experience. For instance, having occupied different positions with respect to a given type of situation gives one better reason to trust one's reactions than one has when one has only held an advantaged or disadvantaged position. Third, in general, our reactions are more reliable to the extent that they are formed in light of interaction and communal reflection with other evaluating agents.

With these general observations about value and emotion in place, I turn to a discussion of empathy. The goal is to clarify the notion of empathy with which I am working, and then to explain its operations. In the final section we will mobilize the foregoing considerations about the relation between emotion and value to explore empathy's role in evaluative inquiry.

II. WHAT IS EMPATHY?

Though "empathy" is a relatively recent term in English, it has already become ambiguous. One source of this ambiguity is its unclear relation to "sympathy"—itself a term with numerous meanings.\(^{19}\) "Empathy" entered our language early in this century as

---

\(^{19}\) A helpful, though tendentious, discussion of the relation between empathy and sympathy can be found in Douglas Chismar, *Empathy and Sympathy: The Important Difference*, in 22 J. VALUE INQUIRY 257, 257-66 (1988). On that account, to empathize is to respond to another's perceived emotional state by experiencing feelings of a similar sort. Sympathy includes empathizing, but entails also having a positive regard or a non-fleeting concern for the
That word was used in German aesthetics to refer to a postulated kind of response to art, in which one first engages in some involuntary bodily mimicry of the work, then projects onto it an emotional response that somehow fits with one's acquired bodily posture. Theodor Lipps eventually came to think the phenomenon could occur in interpersonal cases as well, when people come through unconscious mimicry of another's behavior to feel some emotional response that they then project onto the observed subject. This is the basis for the most familiar contemporary sense of the term "empathy" (a sense we will presently seek to sharpen), on which to empathize is to react to the perceived feelings of another with vicarious emotional reactions of one's own, and empathy is the capacity for, or the occurrence of, such a vicarious experience.

"Sympathy" is a much older term, and has a number of rather different meanings. It was widely used by eighteenth-century philosophers such as Frances Hutcheson, David Hume, and Adam Smith (the so-called "British Moralists") to refer both to mechanisms of empathy (including some like those considered by Lipps) and to various feelings and benevolent motives that those mechanisms can produce. This ambiguity persists in contemporary use. In order to distinguish between what are clearly distinct phenomena, I will use "sympathy" here in a more restrictive sense as the name for the kind of sentiment that responds to perceived harms or threats to another person with concern for that person, and involves some degree of motivation to aid the person. In effect, I am treating sympathy as another name for pity, though the term lacks the connotations of condescension that are frequently associated with "pity." A fuller treatment would have to take account of the wide variety of other phenomena that are sometimes called "sympathetic." On my way of

20. According to Steven Darwall, Empathy, Sympathy, Care, 89 PHIL. STUD. 261 (1998), the first usage of 'empathy' in our language was in 1909 when Edward Titchener coined it to translate Theodor Lipps's use of Einfühlung.


23. This view of sympathy is largely in line with those of Douglas Chismar, supra note 19, and Steven Darwall, supra note 20, at 262-63, 272-79. Both of those authors treat sympathy in more detail than I can here.
using the terms, sympathy is not only not the same thing as empathy, it is not the same kind of thing: sympathy is an emotion, empathy is a way of acquiring an emotion.24

Granting the distinction, though, is there any reason to suppose that empathy will always lead to sympathy?25 If empathy is understood, as above, as a vicarious experience of an observed subject's emotional reactions, then presumably the answer is no—or not directly, in any case. Sympathy is essentially a third-personal feeling, so it makes no sense to suppose that the observer vicariously experiences sympathy that the distressed person (the model) is feeling. The model isn't feeling sympathy, but sadness, or fear, or distress, etc. Still, even if empathy does not directly induce sympathy, it certainly can set the stage for sympathy. After acquiring vicarious sadness through empathy, one is primed to feel sympathy for a bereaved person toward whom one was initially unsympathetic. Indeed, there is surely some plausibility to the thought that any vicarious emotional experience, taken up as if from the perspective of another, makes one somewhat more likely to be interested in the other person—though it is a further question whether this will lead to concern for that person's well-being. Michael Stocker and Stephen Darwall both seem to deny this, going so far as to claim that empathy or empathic understanding of another person is consistent with the detachment of scientific observation ("or even the cruelty of sadism," Darwall adds).26 They claim that whether empathy leads to sympathy depends upon why one is interested in the other person. No doubt this will be a relevant factor, and no doubt it is possible to empathize without ever coming to sympathize. But it is worth noting that if empathy entails (or even typically involves) feeling another's emotional reaction vicariously, "empathic understanding" will never (or not typically) be detached or indifferent in the way that pure observation is.27

24. Cf. Darwall, supra note 20, at 263 (1998) ("Sympathy for a person and her plight is felt as from the third-person perspective of one-caring, whereas empathy involves something like a sharing of the other's mental states, frequently, as from her standpoint.").

25. An excellent discussion of recent literature concerning the relation between certain forms of empathy and altruistic motivation can be found in Shaun Nichols, Mindreading and the Core Architecture of Moral Psychology (unpublished manuscript, on file with author).

26. See STOCKER & HEGEMAN, supra note 2, at 214-217; Darwall, supra note 20, at 261.

27. See Martha Nussbaum, Equity and Mercy, 22 PHIL. & PUB. AFF. 85 (1993), for an interesting discussion of the relationship between full appreciation of another person's circumstances and sympathy or concern for that person.
Now to attempt the promised sharpening of the concept of empathy. My primary concern is with *mechanisms* of empathy, which I define as mechanisms that tend to influence the emotional reactions of one person—the "observer"—so as to or produce a match (roughly, some sort of congruence) between these emotions and those of another person—the "model." To say that someone is empathizing with someone, then, is to say that she is being influenced by some such mechanism. Whether there are such mechanisms, and if so, how they operate, are matters for empirical investigation. I discuss the results of some relevant research in Section three. In any case, all discussion of empathy in what follows will be focused on empathy that is postulated to arise from some such mechanism. Several features of this approach should be explained and defended.

My definition focuses on *mechanisms* of empathy in order to highlight the point that an actual match is neither necessary nor sufficient for the occurrence of the phenomena that have generally been of interest to researchers who have taken themselves to be studying "empathy." It is not sufficient because an emotional match between model and observer might occur by chance, without the operations of any psychological process that even tends to bring about such matches. My account avoids calling such accidents cases of empathy.28 If we know that Bob and Carolyn are each angry about the actions of Congress, and that Bob saw Carolyn getting angry before he got angry himself, we surely need to know more to establish whether Bob is empathizing with Carolyn's anger or getting angry about the same thing independently. So a match, even a match preceded by observation of another party's reaction, is not sufficient for empathy.

A match is also not necessary for empathy on the proffered account. This is in order to countenance the possibility that genuine instances of the psychological processes we are exploring may fail, on any given occasion, to produce a "match" of the relevant sort. Suppose Melanie finds herself feeling sad in the presence of Mark (whose brother has recently died) and that her feeling sad is a result

28. It might be thought that this problem of accidental matches is better treated by adding a causal requirement. One could say that empathy occurs whenever a model's emotional reaction induces a matching emotional reaction in the observer. But if this conception were adopted, then simulation or perspective taking would turn out not to be empathy. This would be an intolerably revisionary result, in light of the paradigmatic role that simulation has played in the scientific study of empathy. Simulation will receive considerable discussion in what follows.
of psychological mechanisms that are responsive to Mark's expressive behavior and that tend to bring about matches between the model and the observer. Suppose, though, that (for reasons we need not specify) Mark is actually not sad about the death of his brother, but relieved. Is Melanie empathizing with Mark or not? Ordinary usage may not be sufficiently determinate to favor either answer; but in sharpening the concept, I want to emerge with a conception of empathy that is informed by empirical study and scientific theorizing as well as ordinary usage. On my account we will say that this is a case of empathy, albeit misfiring empathy. The benefits of this approach in accommodating the science of empathy will become evident in the next section, which surveys recent psychological accounts of empathic mechanisms.

The relation of a "match" between emotional states is intended to be broader than the more traditional suggestion that empathy produces the same emotion in the observer as is experienced in the model.29 One reason for preferring this wider notion is that, on a number of views, the emotional experiences of an empathizer are not actually genuine instances of the same emotion as the model is experiencing—they are imaginary emotions, or quasi-emotions.30 Thus it has been claimed that you are not literally afraid when you empathize with the plight of the person you see in grave danger (perhaps because to be afraid is to feel endangered and/or inclined to fight or flee, and you don't have these attitudes). But even if the emotion you feel is different than the one the model feels, there may be important points of correspondence between them. For instance, it may be that both are unpleasant experiences focused on some source of threat to the model, and both involve viewing that source in a negative light precisely as a threat to the model. There may even be some correspondence in physiological symptoms between these states. Then, I will want to say, we have a match in the (admittedly vague) sense I am using.31

29. Here I follow Barnett and allied theorists who require only that the emotion in the observer be "congruent with" that in the model, rather than identical. See, e.g., Mark Barnett, Empathy and Related Responses in Children, in EMPATHY AND ITS DEVELOPMENT 157, 157-62 (Nancy Eisenberg & Janet Strayer eds., 1987) (and literature cited there). The remarks here about the notion of a match might help to flesh out the relevant sense of 'congruence.'


31. Daniel McIntosh et al., Socially Induced Affect, in LEARNING, REMEMBERING, BELIEVING: ENHANCING HUMAN PERFORMANCE 251, 252-53 (Daniel Druckman & Robert A. Bjork eds., 1994), distinguishes between "concordant" and "discordant" cases in which an observer's emotional experience is "induced by" a model's observable emotions or feelings.
Finally, my definition appeals to mechanisms that tend to influence one person's emotions so as to produce a match between them and another person's emotions. This "so as" is ambiguous in a way that invites us to ask, but does not prejudge, the question of whether it is the function of the mechanisms in question to produce such matches, or whether that is merely their effect. Whether the capacity for empathy has a selective history of the sort that might make it an adaptation is an interesting question. Further discussion will demonstrate some ways in which empathy might be important for our capacity to understand and predict each other's behavior—a capacity which is obviously important to our social lives, and hence to our fitness. And many psychologists believe that there are a host of different mechanisms underlying empathy, including some redundant systems. This superabundance of mechanisms provides some reason for thinking that the capacity for empathy has adaptive significance: where several complex, costly systems produce the same effect, and that effect has an important impact on the organism's fitness, there is a prima facie reason to posit evolutionary function. It seems plausible to me that some empathic mechanisms are adaptations, but I will not attempt to argue for that claim here, since none of what follows depends upon this suggestion.

III. MECHANISMS OF EMPATHY

The richest historical discussions of the mechanisms of empathy are David Hume's and Adam Smith's treatments of what they called "sympathy." These authors had rather different ideas about its operation. In the last thirty years, developmental and social psychologists have been exploring the mechanisms of empathy in considerable detail, and the reality of the phenomenon is by now well documented in a number of experiments. Concordant cases are those where the model and the observer experience affect with the same valence (positive or negative), and discordant cases are those where the valences are opposed. I intend the notion of a match to require more than just concordance, including, at least, a common intentional focus in the two parties' emotions. I leave deliberately unsettled whether socially induced cases of discordant affect could ever be counted as a match, and hence whether such cases will ever properly be called instances of "empathy." That question, it seems to me, may be best decided by the fruitfulness of the different possible extensions of the term "empathy" within psychological research programs. Accordingly, the acknowledged vagueness in the definition of a "match" is principled, and not careless.

32. For a review of such research, see id. at 254; Bert S. Moore, The Origins and Development of Empathy, 14 MOTIVATION & EMOTION 75, 75-79 (1990).
contemporary research into specific mechanisms of empathy has led psychologists toward two distinct paradigms, *contagion* and *simulation*, suggested respectively by the work of Hume and Smith. In this Section, I explain and draw some distinctions between these two forms of empathy.

"'The human countenance'" says Hume33 (quoting approvingly from Horace), "'borrows smiles or tears from the human countenance.' Reduce a person to solitude, and he loses all enjoyment... because the movements of his heart are not forwarded by correspondent movements in his fellow creatures." This is one of the things Hume meant by "sympathy," the curious empathic phenomenon in which an emotional state is transmitted from model to observer. It is a "propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments, however different from, or even contrary to, our own."

This aspect of sympathy is important for Hume's account of moral motivation, since he thinks it serves as a partial explanation of the other kind of sympathy: benevolent concern for the well-being of others, aroused through vicarious experiences of their pleasures and pains.36

The quote from Horace, and some of the other remarks Hume makes about empathy (as we will anachronistically call it), suggest a view of empathy as an involuntary "catching" of another's emotion, induced somehow by the model's expressive behavior. Recent scholarship refers to this sort of process as "contagion." What is striking about this kind of empathy, which the epidemiological metaphor highlights, is the apparently noncognitive nature of the mechanism. It seems that without any reflection on what state the model is in, indeed without conscious awareness that he is attending to the model's state at all, an observer can find himself feeling an emotion that matches the model's. Following these strands in

33. DAVID HUME, AN ENQUIRY CONCERNING THE PRINCIPLES OF MORALS 54 (Open Court Pub. Co. 1930) (1777).

34. DAVID HUME, A TREATISE OF HUMAN NATURE 316 (L. A. Selby-Bigge ed., 1896).

35. Contagion is only one of the things that Hume treats under the rubric of sympathy. In other contexts, the term refers to other aspects of our "natural" tendency to respond emotionally to others. Recent discussions of Hume on sympathy have tended to focus on these other aspects. See generally ANNETTE BAIER, A PROGRESS OF SENTIMENTS (1991).

36. Though Hume expressly disavows the search for any explanatory principles more basic than the tendency of persons to take pleasure and pain at the pleasure and pain of others, he also thinks we can infer a concern with the causes of emotional expressions (i.e., with good and bad behavior) from the fact that we are moved by the expressions themselves. See HUME, supra note 33, at 54 n.1.
Hume's thought, we will treat him as holding a noncognitive, contagion theory of empathy (though other strands pull in other directions). Recent research on empathy not only confirms the existence of such contagion, but traces its causes to just the mechanism that Hume and Horace suggest—an unreflective transfer of emotional expressions from model to observer.

In fact, the best understood mechanism of contagion seems to run through mimicry, particularly mimicry of facial expression. Mimicry is a phylogenetically ancient phenomenon, playing an important developmental and social role in many animal species, including humans. Early in development, infants engage in mimicry. One familiar example is contagious crying in a nursery. Among human adults, facial mimicry in particular is extremely common, and is often involuntary. Dimberg showed observers slides of posed angry and happy faces. Using electromyographs to measure muscular activity, he found that activity of the zygomatic muscle (used to pull back the cheeks when smiling) was high in observers of happy faces, and that activity of the corrugator muscle (used to wrinkle the brows) was high in observers of angry faces. Berger and Hadley recorded muscular activity in the arms and lips of observers watching arm-wrestling and stuttering videotapes, finding higher levels of activity in the observers' muscles corresponding to the muscles being used by the models. On a lighter note, Zajonc and his colleagues found empirical support for the folk adage that old couples look alike, and they argue that this is best explained by repeated mimicry of each other's facial expressions over many years. Other studies confirm the general phenomenon: consciously or unconsciously, observers

37. At various points Hume suggests a more cognitive picture, according to which the observer forms an idea of the emotion the model is experiencing as a result of observing her expressions, and it is this idea of the model's emotion that somehow causes the emotion itself to occur in the observer. See Robert Gordon, Sympathy, Simulation, and the Impartial Spectator, in ETHICS 727, 727-29 (1995). Gordon points out some of the difficulties with this picture. One such difficulty is that requiring that contagion be mediated through beliefs makes it hard to explain how infants could acquire emotions contagiously without attributing implausibly rich cognitive capacities to them. See id.

38. See Janet Strayer, Children's Concordant Emotions and Cognitions in Response to Observed Emotions, 64 CHILD DEV. 188, 188-201(1993).


41. R. B. Zajonc et al., Convergence in the Physical Appearance of Spouses, 11 MOTIVATION & EMOTION 335, 335 (1987).
frequently mimic the behavior of models—particularly their expressive facial behavior.\textsuperscript{42}

Of course, that observers mimic the facial gestures of models does not by itself show that observers experience the emotions that are "expressed" by these gestures in the model. But there is substantial evidence that they do. Other studies show that electrodermal responses, changes in heart rate, and changes in self-reported feelings are all correlated with models' expressive facial behavior.\textsuperscript{43} Furthermore, these correlations are not independent of mimicry. Vaughan, Lanzetta, Bush, and others found that observers asked to inhibit their own facial activity experienced less empathic affect, and had diminished changes in heart rate and electrodermal activity compared with those who were given no instructions. Vaughan and Lanzetta also found that observers asked to adopt facial expressions corresponding to those of the model experienced greater changes in skin conductance and heart rates.

How can this be? How can facial gestures we think of as expressions of underlying emotional states actually cause the very states we take them to express?\textsuperscript{44} Recent studies suggest that it may be a product of "feedback," the phenomenon whereby facial and other gestures cause internal changes that constitute or themselves cause emotions. Eschewing the term "expression" for its obvious presuppositions about the direction of causal influence, Robert Zajonc and some others have begun to investigate facial "efference": facial gestures which produce internal changes through a kind of feedback. Ekman instructed actors to form their muscles, one by one, into six "basic" emotional expressions. They reported dramatic changes in heart rate varying with the nature of the actors' facial movements, even though the actors' task was not described to them in terms of emotional expressions. Lest it be thought nonetheless that the actors performed the facial task by attempting to call up particular emotions, Zajonc, Murphy and Inglehart performed a similar experiment, putatively about phonetic research. They instructed

\textsuperscript{42} A useful survey is described in McIntosh et al., supra note 31.


\textsuperscript{44} Carl Georg Lange and William James notoriously held that the bodily disturbances commonly thought to be effects of emotions were in fact their causes. CARL GEORG LANGE & WILLIAM JAMES, THE EMOTIONS 5-6, 13 (1922). But that view hasn't attracted many contemporary adherents. For an astute critical discussion, see ROBERT GORDON, THE STRUCTURE OF EMOTION 86-88 (1987).
subjects to make particular vowel sounds, the making of which in fact requires facial configurations akin to smiling and frowning. Subjects were found (based on self-reports) to prefer feelings generated by making the vowel sounds requiring a smiling expression to those generated by making the sounds requiring the frowning one. Additional studies by these and other researchers contribute to a large body of data that demonstrates that facial efference can both initiate and modulate emotional experience.45 While the details and explanation of these phenomena remain controversial,46 it seems clear that the combination of (facial) mimicry and feedback is one mechanism by which emotional contagion occurs.

One possible shortcoming of contagion is a limitation on the range of emotional experiences whose transmission it might plausibly explain. Hume invoked it primarily to describe how pleasures or pains in a model might be transferred to an observer. The recent research just canvassed extends the theory to cover transmission of any emotions with distinctive expressive physiologies through feedback. But these “basic emotions” still represent a fairly narrow span of emotional states: anger, fear, sadness, happiness, surprise, and disgust, according to the best-developed research program on basic emotions.47 For those who wish to make more fine-grained distinctions among types of emotion by appealing to different sorts of beliefs and commitments that these emotions involve, contagion may appear an impoverished account of empathic transmission. Thus, for instance, it may be that contagion is incapable of distinguishing


46. The only functional explanation I am aware of for the facial feedback phenomenon is Zajonc's updated version of a theory originally proposed by the turn-of-the-century physician Israel Waynbaum, who suggested that facial gestures permit regulation of blood flow to the brain. See R. B. Zajonc, Emotion and Facial Efference: A Theory Reclaimed, 228 SCIENCE 15, 15-21 (1985); cf., DEAN FALK, BRAINDANCE 155-59 (1992). Zajonc argues that this vascular theory is largely correct, though the true function of these facial movements is control of brain blood temperature, rather than blood supply. See Zajonc, supra, at 15-21. He makes a compelling case, the truth of which I feel ill equipped to assess. It is worth noting, though, that the enormous importance of maintaining appropriate brain temperature has been the focus of at least one recent account of the extraordinary evolution of the hominid brain. See id.

47. See PAUL EKMAN, THE FACE OF MAN: EXPRESSIONS OF UNIVERSAL EMOTIONS IN A NEW GUINEA VILLAGE (1980). I suspect that further research will ultimately expand the list of pan-cultural emotional syndromes to include some (such as shame) that are now typically regarded as culturally idiosyncratic. But it should be granted that ordinary thought, with its cognitive sharpenings of various emotional states, has far more emotional categories than our bodies have distinctive emotional phenotypes.
between (i.e., transmitting just the appropriate one of) anger, moral indignation, and social offense.

If the above difficulty for contagion is correctly diagnosed as arising from cognitive distinctions among different kinds of emotional reaction, the solution might be a more cognitive account of empathy. Adam Smith developed a theory of sympathy that made room for a far greater repertoire of empathetic feelings. He also held that “we can form no idea of the manner in which [other men] are affected, but by conceiving what we ourselves should feel in the like situation.” It is through imaginatively projecting ourselves into another’s position, and then introspecting to see how those circumstances make us feel, that we gain any empathic appreciation of the feelings of others. This is the simulation theory of empathy. Much of the recent psychological work on empathy has taken simulation (or “perspective-taking,” or “role-taking” as it is sometimes called) as the paradigmatic empathic process.48

The simulation theory has recently been defended by several influential philosophers and psychologists as a general account of our capacity to attribute mental states to others, and to predict their behavior.49 The general thought is that, in order to work out what someone else believes, desires, or feels, and hence, ultimately, what she is likely to do, we imaginatively enter into her position, and then see what we find ourselves imaginatively believing, desiring, or feeling. That is, we run our own belief-forming (or desire-forming, or emotion-forming) processes “off-line,” using imaginary inputs that appropriately resemble the circumstances of our model, then we attribute to the model the real counterpart to whatever imaginary state emerges from this simulation in us.50 While some details vary between accounts, contemporary simulationists are united in


50. Smith’s reference to “conceiving what we ourselves should feel in the like situation” is ambiguous here. In simulating King George’s feelings at the coronation, am I to imagine how I (Justin D’Arms) would feel if crowned King of England, or imagine being him, being crowned? Contemporary simulationists prefer the latter description, which stands a better chance of producing a matching reaction. I, after all, would be astonished to be crowned King of England. George presumably was not.
accepting these broad outlines, and in seeing this account as an alternative to the more traditional "theory theory" of mental state attribution. The "theory theory" holds that the capacity to attribute mental states to other people (and to predict their behavior in light of these attributions) involves nothing more than the application of a (perhaps tacit) theory of mind, or "folk psychology." In particular, such attribution does not require a simulation of the target's circumstances. This theory is offered as an account of mental state attribution, not an account of empathy. I mention it here because the opposition between it and the simulation theory will be useful in clarifying the simulationist's position about empathy.

Like Smith, modern defenders of simulation as an account of empathy have occasionally made an overly extravagant claim for it: that it is the sole effective mechanism of empathic transmission. But a simulationist need not make so strong a claim. Hoffman, for instance, defends a pluralistic view of empathic mechanisms. He has found evidence supporting a model of child development in which the capacity to imaginatively place oneself in another's shoes is a discrete stage in the development of "prosocial" behavior, following after the development of several distinct noncognitive capacities. One significant, if unsurprising, finding is that once infants develop a conception of others as distinct from themselves, the character of their empathic responses becomes more complex. Rather than merely experiencing distress at the presence of distressed models, they begin responding with a range of feelings appropriate to the model's situation. In adult simulations, the range of possible experiences for empathic transmission is presumably broader still, since adults typically possess both a more nuanced repertoire of cognitive-cum-affective responses and a more sophisticated ability to


52. See, e.g., ALBERT BANDURA, PRINCIPLES OF BEHAVIOR MODIFICATION 196-99 (1969); Ezra Stotland, Exploratory Investigations of Empathy, in ADVANCES EXPERIMENTAL SOC. PSYCHOL. (Leonard Berkowitz ed., 1969). Defenders of simulation as a general account of mental state attribution tend to be more ecumenical, holding that simulation is simply one (perhaps especially important) means by which we are able to attribute such states to others.

think our way into another's circumstances. It appears then, that the
cognitive/imaginative operations of simulation permit transmission of
more finely grained emotional responses than does contagion.

To decide whether this appearance is misleading, let's consider
contagion in a little more detail. As described so far, contagion has
an important shortcoming. Its workings may explain the transmission
of an emotion from one party to another, but they do not yet explain
the correspondence in the intentional focus of that emotion, which I
suggested earlier is a requirement for generating a match between the
observer's and model's emotions. If Bob catches Carolyn's fear by
first mimicking her expressions unconsciously and then acquiring fear
by efferent feedback, it is not yet clear how his emotion could come
to match hers in the sense of being fear about the same thing. There
are two sorts of answers to this question. One appeals to the
phenomenon of gaze-tracking.\textsuperscript{54} It is a brute fact about humans and
many other animals that we tend to turn our attention in the direction
that another person is looking.\textsuperscript{55} Thus, Bob may match Carolyn's fear
by first catching some undirected fear (or proto-fear) through
contagion, then looking to see what it is that she's looking at, and
focusing his fear on that. Another possibility is that once he begins to
feel something like fear as a result of contagion, Bob casts around in
his environment for something to which this makes sense to him as a
reaction. That is, he searches for something fearsome. Whether he
alights on the same thing as Carolyn did may depend upon many
factors: including (1) the degree to which the target of her emotion
can be read off her expressive behavior; (2) the number of possibly
salient targets for fear in the environment; and (3) the degree to
which Bob is himself susceptible to fearing the thing that Carolyn
fears.

In order to explain how it can produce a match even of such
basic emotions as anger and fear, then, contagion must be understood
as operating in a rich context that includes the environment in which
the contagion occurs and the observer's beliefs about that
environment. But once this context is properly in view, it is less
obvious that contagious empathy will be restricted to transmitting
basic emotions. It may be that the initial infection, so to speak,

\textsuperscript{54} See Goldman, supra note 49, at 106.

\textsuperscript{55} Although I cannot cite any published research on this point, my own parental
experience establishes that the phenomenon can arise at least as early as six months after birth.
It seems to be as close to "innate" as human behavior gets.
transmits one of a relatively narrow range of basic emotional types of experience. But then, in fixing on an object toward which the experience will be directed, the observer may also be supplying precisely the cognitive elements to the experience that determine whether it is, say, a case of moral indignation or (merely) of social offense. Even so, I grant that there may be a narrower range of experiences that are amenable to contagious transmission than to empathic simulation. My point here is simply that contagion occurring in normal circumstances and in the context of a normal human psychology may well be capable of transmitting even fairly nuanced forms of response.\(^{56}\)

Once we allow mechanisms of contagion to help themselves to all these cognitive resources, though, we may seem to have effaced the difference between contagion and simulation. Certainly there is now some role for cognition in both sorts of process. But the role is not the same. In fact, there remains a difference between the processes that will turn out to be very important. To explain this, we need the notion of "sensibility": a disposition to experience particular kinds of emotional reactions in response to particular sorts of cues. In other words, one's sensibility is one's pattern of tendencies toward emotional reactions. A person's reactions may be more or less consistent, and hence his sensibility may be more or less determinate. It includes such things as his sense of humor, his proclivity for embarrassment, his sensitivity to particular sorts of slights, and so on. The difference between contagious transmission, the emotional experience that the observer acquires depends crucially on the model's sensibility; whereas in a simulation, it is only the observer's own sensibility, and not the model's, that determines what emotion(s) she experiences.

Suppose Angus behaves toward Carolyn in a way that might elicit either offense or amusement, depending upon one's sense of humor.\(^{57}\) Suppose Carolyn is amused. Bob might catch her amusement through contagion, or he might try to think his way into

---

56. It may be more accurate to say that contagion *initiates* the transmission of these cognitively richer responses. Full transmission, by hypothesis, sometimes requires not merely the noncognitive operations of mimicry and feedback, but also the right environmental cues and/or background beliefs in the observer. For convenience, though, I will continue in what follows to use "contagion" more inclusively, to refer to a transmission process that can include these further features.

57. Of course, which sort of response such behavior actually elicits would depend upon other factors as well as Carolyn's sense of humor, but those need not concern us here.
her circumstances, and see how he finds himself feeling. If contagion is the effective mechanism, then it is Carolyn's sensibility (her sense of humor in this case) that begins the process that produces Bob's amusement. He is amused because she is inclined to find Angus's antics funny, not rude, and he is infected with her attitude. Whereas if Bob simulates, it is Bob's sense of humor that explains the amusement. He is amused because he finds Angus's behavior funny when he imagines being the one at whom it is directed.

Why is a simulator restricted to exercising his own sensibility, rather than imaginatively adopting his model's? Contemporary simulation accounts have tended to be vague about how much of the simulator's own psychology stays with him when he imaginatively places himself in someone else's shoes. In order to simulate how the president feels at the news that his wife is moving to New York, do I try to imagine that my wife is moving to New York and see how I feel? Or do I first try to imagine that I'm a philanderer with the weight of the Free World on my shoulders, and then try to imagine that my wife is moving to New York? Of course I stand a better chance of producing a match between my state and the model's (and hence a better chance of accurately predicting his response) the more of his personality and circumstances I am able to take on in the simulative exercise. So contemporary simulationists allow that part of what is taken on in a simulation can be beliefs and desires that the model has, but the simulator doesn't. Crucially, though, if this account is to maintain its autonomy from the "theory theory," simulation must not require that the simulator invokes beliefs about how the model would react to a certain set of stimuli. It is important that what actually determines the state that the simulator gets into through simulation are his own mechanisms of emotion-formation, or belief-formation, or desire-formation—even if these processes are allowed to take as inputs some imaginary attitudes quite different from his own.58 So there is a restriction on just how far we can

58. In replying to the charge that simulation is itself simply an exercise of theory about states that the model ("target") will be in, given certain inputs, Goldman argues that "it seems far-fetched to suppose that my ability to gauge what will amuse you is based on a theory of humor (of what amuses people). I do not possess any general theory of this sort. More plausibly, I gauge your probable reaction to a joke by projecting my own. (There can be adjustments here for factual information about interpersonal differences, but this is just a corrective to the basic tactic of simulation.)" Alvin I. Goldman, Interpretation Psychologized, in FOLK PSYCHOLOGY, supra note 49, at 83, 84. So the simulator must exercise his own sense of humor, it seems. The accuracy of the simulation in arriving at the model's likely response then depends upon the similarity between the simulator's sense of humor and the model's. "A
understand a simulator as going when he "recenters his egocentric
map." While he may take (imagined versions of) another's beliefs
and desires as input, he must bring with him the psychological
processes that generate new states on the basis of those inputs.
Otherwise, this will not be a simulation, but a piece of theorizing
about how the other person would react. Hence, in the case where
the output is an emotional reaction, the simulator must bring his own
sensibility to bear on the model's circumstances.60

It follows from these considerations that the reliability of
simulation in producing matches between the emotional states of the
observer and those of the model depends strongly on the degree to
which they are alike in sensibility.61 Whereas, since contagious
transmission is initiated by a reaction arising from the model's
sensibility, its reliability seems not to depend upon any such
similarity. Admittedly, there is a way in which the observer's
sensibility may be pertinent to the emotion he acquires through
contagion. If, for instance, Bob were so dour that he would never,
himself, be amused by Angus' behavior, then he may be incapable of
contagiously acquiring amusement from Carolyn—or, if he does
acquire it, he may direct it at something other than Angus. If so,
Bob's own sensibility can be said to influence the results of contagion.
Note, though, that emotional reactions can be difficult to suppress,
and the salience of Angus' interaction with Carolyn may make it
difficult to focus his attention on something else as an object of
simulation of some target systems might be accurate... if (1) the process that drives the
simulation is the same as (or relevantly similar to) the process that drives the [target] system,
and (2) the initial states of the simulating agent are the same as, or relevantly similar to, those of
the target system." Id. at 85.

59. See Gordon, supra note 37.

60. This requirement on simulation may appear to be simply a product of a turf division
within the philosophy of mind, rather than a principled restriction on what kinds of empathy are
possible. But there are good independent reasons for thinking that, insofar as simulation is a
way of empathizing, it depends on the observer's sensibility. Consider an attempt to simulate
the reaction of a person whom you know to have a very different sensibility from your own. He
is easily offended, let's suppose, and seldom amused, while you are thick-skinned and quick to
laugh. You may be able to predict his dour reactions to something that would have amused you.
But can you simulate them, and so become empathically irritated at something you are naturally
inclined to find quite funny, and not at all offensive? It seems to me highly doubtful that you
can. The point is not that simulationists are committed to supposing that we can't recognize
differences between our own sensibilities and those of other people. The point is that
simulation is not promising as a device for generating emotional congruence between people
with such differences of sensibility. Whatever its other failings, contagion therefore offers a
prospect that simulation does not. I am grateful to Heidi Feldman for pressing me to clarify this
issue.

amusement. More generally, it is a familiar phenomenon that people who wish to take a stern attitude, and who might be disposed to do so on their own, can be drawn in by a friend's amusement and find themselves laughing at something they would disapprove of under other circumstances. In any case, the main point is that only contagion provides vicarious access to the sensibility of the model. We shall see in the next section how this affords contagion a distinctive role in evaluative inquiry.

Of course, the forgoing discussion does not tend to show that contagion is more reliable than simulation, it merely demonstrates one source of potential error in simulation from which contagion is immune. In fact, each of these processes can go wrong in several different ways, and it is difficult to see how to ground any general claim about which is more reliable. The reliability of simulation seems to depend on several things: how good we are at thinking our way into the other's position; how emotionally responsive we are to imagined circumstances; and how much our own dispositions to respond resemble those of our models. The reliability of contagion depends on rather different factors, including the degree to which the model's expressive behavior accurately signals her emotional state, the observer's ability to discern that behavior, the reliability of mimicry and feedback, and the range of factors discussed above that influence what the observer focuses her acquired emotional reaction upon. In what follows I assume that each of these processes are sometimes reliable, and explore the consequences of this assumption for the epistemology of value.

IV. EMPATHY AND DISCOVERY

I suggested earlier that emotions play a significant role in evaluative thought, in part because many emotional reactions present themselves as sensitivities to values in the world. Deciding what has

62. See Roy A. Sorensen, Self-Strengthening Empathy, 58 PHIL. & PHENOMENOLOGICAL RES. 75, 75, 77 (1998) (giving an interesting argument for the conclusion that evolutionary pressures may have operated to make it the case that our dispositions to react are likely to be quite similar). The argument runs roughly as follows: Average people are more psychologically similar to other people than nonaverage people are. Being psychologically similar to others makes one a better simulator of their mental states. Being a good simulator has important benefits for fitness since (1) one's ability to anticipate the behavior of others is crucial to success in various ways and (2) behavior can be anticipated only if one can make good predictions about the mental states of others. Hence, selection pressures favor average psychologies; over time the population should be expected to have tended towards increasing psychological similarity.
value often depends upon having such reactions, and then reflecting on their appropriateness. I highlighted a “positive” role for these emotional experiences: they sometimes make us aware of values that we had not previously recognized. In other words, some emotional episodes provide us with input that can expand our sensibilities and convictions so as to make room for new sources of value and interest. We are now in a position to see how empathy can assist in this process.

Let’s begin with a mundane simulation. Suppose a group of children are in the habit of teasing another child. One of the teasers misses school one day, and the next day his friends gleefully report to him about the especially good fun they had teasing the victim yesterday. Now suppose that, for some reason, on hearing these reports our protagonist begins to imagine what it would be like to be the object of jibes like the ones his friends were administering. He might find himself feeling pretty bad as a result of this simulation, and thereby come to think that what they did was cruel. If so, we could say his empathy produced an emotional reaction that led him to a new evaluative take on the circumstances.

Does our newly sensitized schoolboy have any reason to think he has learned something about value as a result of this simulation, and not simply been led into weakness by these feelings? Perhaps he does, though of course he may or may not realize this. Given the first assumption articulated earlier, he should accord a defeasible presumption of warrant to his reaction. In other words, the fact that he found himself feeling hurt and indignant in his simulation counts as some reason for thinking the teasing was cruel. Of course the presumption is defeasible, as always, and the mere fact that he responded in this way is only one consideration among many. But the point is that simulating the reactions of another person does seem to provide potentially relevant data to be weighed in thinking about this sort of conduct.

In fact, simulation is a familiar and epistemically respectable way to generate the emotional responses that influence our evaluative judgments. After all, we quite commonly think that it is appropriate to suspend judgment about something until we have had a chance to consider how things might feel if we were in a different position. We allow that circumstances look different from different points of view. It is surely only reasonable to grant that judgments we make after
seeing what things feel like from several such points of view are ipso facto better judgments. This line of thinking accords with the general principle broached earlier, that reactions are generally more trustworthy when informed by greater experience—for instance, by experience in both advantaged and disadvantaged roles.

What about contagion? I claim that contagion, too, has an important role to play in evaluative inquiry. Since contagion transmits emotional responses from model to observer, it can enable the observer to see her circumstances in emotionally laden ways that sensitize her to evaluative features she had not previously recognized. But contagiously acquired emotions do not enjoy the same intuitive support as simulated ones, in part because contagion does not arise directly out of reflection on other ways of considering the circumstances.

Consider an example. Suppose Mara is angry at Tracy for canceling dinner plans with her and Sam at the last minute in order to stay late at work. Sam’s a more understanding sort. He’s generally slow to conclude that he’s been slighted. Indeed, his friends sometimes tell him that he lets people treat him like a doormat. Sam initially sees Tracy’s late cancellation as a sign of her professional dedication, a trait he admires. But as Mara sits glowering at the restaurant, Sam finds his relaxed mood giving way to irritation, too. Mara points out that Tracy is always doing things like this. She just shouldn’t make plans on weeknights if she won’t keep them—but she wants to have something to do in case she runs out of steam at work. Of course, the irritation Sam is feeling might go either way. He could become annoyed with Mara for being so crabby. But her observations about Tracy’s selfishness might instead combine with Sam’s feelings to bring him around to Mara’s view of things: that Tracy has mistreated them. To the extent that mistreatment is tied to feelings of anger, if Sam thinks these feelings of aggrievement appropriate here, he will think Tracy treated him badly. Empathy will have played an important role in bringing him to this view of the situation.

The idea that a contagiously induced emotion can help one to see previously unrecognized evaluative features of ones circumstances is

63. Of course, additional information can sometimes mislead, and in this respect perspectival information is no different. But that’s just to say that the relationship between information and good judgment is not monotonic, which is no reason for doubting that, in general, getting more information improves judgment.
even more plausible with respect to values that are response-
dependent in the sense articulated earlier. Because response-
dependent concepts accord sentiments a central role in questions 
about their very application, sentimental experience and reflection on 
its appropriateness have a special relevance to evaluation in response-
dependent terms. If (on some other occasion) Sam is contagiously 
affected by Mara’s amusement, he sees features of the circumstance 
as funny. From such a position he is especially likely to come to share 
Mara’s evaluative stance: that the situation is funny. When the 
question of whether some evaluative concept applies is a matter of 
sentiment, a matter of which feelings are appropriate, it becomes 
possible for the impinging emotional responses of another person to 
arrive as invitations (with sketchy directions) to see matters as she 
does. This is because when we have an affective response to a 
situation, we look for, and are sensitized to, features of the situation 
that render that response at least intelligible, and at best appropriate. 
Furthermore, when the values in question are response-dependent 
concepts, concluding that the response is appropriate is concluding 
that the thing has the value in question. This was the point of the 
second assumption articulated earlier. In general, the tighter the tie is 
between adopting an evaluative conviction and having some 
emotional response, the more useful empathy will be in inducing 
changes in the observer’s evaluative stance.

Still, there is an apparent problem about contagion that the 
restaurant story may help to illustrate. For surely the pretense that 
we can regard convergence in emotional experience as any kind of 
reason for thinking that we are coming to recognize what’s really 
shameful, funny, offensive, and so on, must be abandoned once we 
appreciate the blind, noncognitive, and mechanistic nature of 
contagious empathic transmission. To the extent that Sam is brought 
around to sharing Mara’s view of Tracy’s behavior by unwittingly 
“catching” her anger, he (and we) may be inclined to doubt whether 
Tracy really mistreated them after all.

Did Tracy mistreat them? Maybe she did, maybe she didn’t. As 
always with such questions about the appropriateness of emotional 
reactions to a situation, one must consider the details of the situation 
and its history—and I haven’t written these into the story. My claim 
is simply that there are a number of reasons for thinking that Sam’s 
contagiously induced anger can properly play a role in his 
deliberations about this question. First, again, this is an emotion. As 
such, it is entitled to a defeasible presumption of warrant. What
makes this case unusual is that the emotion is a product of someone else's sensibility, rather than of Sam's.64 Perhaps Sam has special reasons to be suspicious of his anger on just these grounds, if he has reason for thinking Mara's sensibility is problematic (that she's generally irascible, say, or jealous of Tracy). But I would urge that Sam has no grounds for systematically ignoring contagiously induced emotions because he has no grounds for supposing that other people's reactions are generally inappropriate.

The sort of surrender to another's sensibilities that contagion requires of us can be unsettling, to be sure. But do we have special reason to be suspicious of the ways in which it influences our judgment? To the extent that we take seriously the idea that there are warranted and unwarranted ways of responding to our world, that feelings can be appropriate or misguided,65 we must acknowledge the possibility that other sensibilities can be better than our own—better at generating appropriate feelings in response to the situations that confront us. Contagiously induced affect has no special claim to our allegiance, of course. As with all our feelings, we can assess its appropriateness and find it wanting. But to deny it any role in our reflection is no more justifiable than would be a generic refusal to consider the normative opinions of others when they conflict with our own.66

The psychological evidence suggests that both simulation and contagion influence our feelings. The question now is whether this is good news or bad. Should we celebrate only simulation, and regard contagion as a form of contamination? My answer is no, because contagion offers a prospect that simulation does not. In a simulation, the simulator brings her own sensibilities to bear on another person's situation. If this produces new feelings in her, it is her own sensibility

64. This is on the assumption that contagion has succeeded in this case in producing a match between Sam's emotion and Mara's. David Merli reminded me that contagion can misfire, and when it does it may induce an emotion that comes, in a sense, without the backing of anyone's sensibility. It is an interesting question whether such emotions should also be accorded a presumption of warrant. This would depend on articulating a fuller rationale for the "first assumption." For present purposes, we are assuming the contagion is reliable.

65. Sentimentalists, of course, will want to take this idea very seriously, since it is with such talk that sentiment theory hopes to capture the aspirations toward objectivity that are such a crucial part of ethical and evaluative discourse.

66. In this connection, there are affinities between my project and some feminist critiques of moral theory. Not only have feminist theorists urged the importance of a variety of emotional responses to understanding morality, they have also called into question the conception of moral agency as isolation and distance from others. Cf. BAIER, supra note 3; GILLIGAN, supra note 3; SCHEMAN, supra note 3.
that explains the presence of these feelings. The model's position only determines the nature of the inputs on which the simulator exercises her sensibilities. Contagion does something more radical. It gives the observer vicarious access to the sensibilities of another. If Mara's anger produces anger in Sam through contagion, he may be enabled to see features of the circumstances and reasons for action that he would never have experienced through mere simulation, since his sensibility is very different from hers. Through the lens of an emotional experience he has borrowed from her, Sam is presented with an alternative way of viewing the facts. He sees Tracy as mistreating him, as taking advantage of him—whereas, if left to his own devices, Sam might never have seen her behavior this way. If this leads to a change in Sam's judgment, that will be a change that arises not simply from exercising his own sensibility on new inputs, but from being influenced by the deliverances of another sensibility.

Such vicarious experiences can influence not only our convictions about particular cases, but also our own patterns of reaction. One's sense of humor, for instance, can be expanded to include new sources of amusement. As a result of being brought to see something funny in a friend's antics through contagious amusement, one can come to see that funniness more generally in her company, where before one was inclined to find that part of her character something to be tolerated rather than reveled in.

To return to the restaurant episode, Sam's vicariously aroused anger at Tracy may produce or at least initiate a change in the way he relates to others, by leading him to expect better treatment from them than he has in the past. If so, contagion has played an important role in changing his sensibility. Furthermore, in the case at hand, the change in question may plausibly be regarded as an improvement. It is important to recognize that this is a change simulation would not have affected. If Sam had not been present at the canceled dinner, and had merely been imagining what it was like to be stood up in this way, we can predict that he would not have found himself angry as a result of that simulation because his own sensibility was not equipped to construe the circumstances that way. By providing him with vicarious access to a new way of taking Tracy's behavior, contagion has done something for his evaluative perspective that simulation could not.

Of course, not all changes in a person's sensibility are changes for the better. One can be drawn into hateful and pernicious forms of reaction as well, and these too can alter one's sensibility in
problematic ways.\textsuperscript{67} We should not be overly sanguine about the influence of contagion in this connection. But remember the earlier principle that reactions are in general more reliable to the extent that they have been subjected to wider influence. It seems to me that, as a general matter, we should welcome contagion because of the way it challenges an all-too-familiar kind of evaluative parochiality or complacency. Think of the task of evaluative inquiry as that of trying to move beyond one's present evaluative perspective, to one which is more fully responsive to the values of things. Given this conception of the task, one must be struck by the scarcity of resources that might help us transcend the perspective from which we initiate our inquiry. I've tried to show how contagion provides some such resources by offering a point of contact and interchange between our own sensibilities and those of other people. Perhaps that is enough to justify the risks it brings. Philosophical critics of a sharp distinction between passion and reason,\textsuperscript{68} and psychological studies of the role of affect in cognition\textsuperscript{69} have converged in recent years in agreement that our emotional responses influence our judgments, both evaluative and even apparently descriptive. We should reject the pernicious idea that any such role played by the passions in the formation of our judgments must be a form of contamination because it would leave us with a conception of reasoned judgment-formation that human beings could not live up to. Furthermore, it would leave us utterly without a foundation for judgments deploying the plainly anthropocentric, response-dependent concepts that play such an important role in our lives. But if we do reject the pernicious idea, allowing instead that our emotional responses can sometimes aid our judgment, then we need to understand the influences of empathy upon these responses. If I am right, these influences are important for advancing evaluative inquiry.

I am no expert on the law. But I hold out some hope that this Article offers a contribution to the discussions of experts, since it seems to a layman that the claims I have been making are relevant to

\textsuperscript{67} It may be that, in some cases, even inappropriate and pernicious emotional reactions are good for us, in some ways. At a minimum, they can teach us a sobering lesson about our own capacities. And discovering how the other side feels is a way of learning to understand how the other side thinks. This can be worth knowing in its own right, and may also assist in the difficult task of articulating defenses of our preferred evaluations that will seem relevant to people with different and disturbing presuppositions.

\textsuperscript{68} See, e.g., DE SOUSA, supra note 1; GREENSPAN, supra note 1.

\textsuperscript{69} See generally EMOTIONS, COGNITION, AND BEHAVIOR, supra note 53.
a wide variety of legal processes and legal reasoning. Judges, juries and legislators are routinely called upon to make evaluative judgments of many kinds. If my claims here are plausible, then those who study these forms of judgment cannot afford to ignore their intimate connections to emotional responses. Furthermore, in light of the social character of legislative and judicial proceedings, the emotional influences in question will predictably be subject to many forms of empathetic process. Of course, legal deliberations have a narrower focus than does generic evaluative inquiry, and it may be that empathy's role is more salutary for the latter enterprise than the former. Be that as it may, there should be no doubt that empathy is at work in legal deliberations as well, and any adequate account of these phenomena will need to understand its role.

70. I do not mean to suggest that these influences have been systematically ignored. Indeed, the existence of this Symposium is a testament to the increasing recognition of the importance of emotion to legal scholarship.