Moral Psychology

Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development

edited by Walter Sinnott-Armstrong

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

5.4 The Makings of a Moral Sensibility: Replies to Commentaries

Victoria McGeer

I am grateful to my commentators for their thoughtful responses to the speculative ideas explored in my chapter. These ideas are largely speculative because, despite a recent surge of interest in atypical moral psychology, it remains a largely uncharted area of interdisciplinary research. Hence, there are very few studies on which to base solid conclusions and very many questions—both empirical and conceptual—still left to answer. Nevertheless, what makes even the modest body of research in this area so tantalizing is the difficult issues it raises on two separate but related fronts. The first is more general, relating to long-standing philosophical debates about the nature of moral judgment and moral motivation. The second is more particular, relating to the specific difficulties involved in investigating atypical cognitive-affective profiles such as those found in autism and psychopathy. I say these two sets of issues are interrelated because our sense of what it is to have a moral sensibility is very much shaped by our understanding of the so-called normal case, sometimes making unusual departures from this norm quite difficult to characterize. In consequence, certain disagreements—for instance, about whether autistic individuals have a genuine but "impaired" moral sensibility—may not in the end turn on facts about specific cognitive and/or affective capacities, but on whether such capacities, and the behavior they motivate, constitute a genuine variety of moral agency. This brings us back to more general philosophical debates about the nature of moral judgment and moral motivation.

In this context, it seems fitting to ask about what we really gain philosophically by studying atypical moral psychology. After all, if we could simply take the presence or absence of a moral sensibility as (detectably) given, then it would make sense to investigate what cognitive and/or affective capacities are "spared" or "impaired" in particular disorders so as to determine what grounds this sensibility. Maybe this would even go some way toward settling the philosophical debate between sentimentalists and

rationalists. Such has been the presumption, at any rate, on the basis of which many of the discussions about psychopathy and autism—including my own—have proceeded. However, in working on my own contribution to this volume, I came to realize that the issues raised by these populations are much more interesting and complex than this straightforward argumentative strategy suggests. Thus, I was led by degrees into a more complicated exploration of what in particular could be going on in autism, and also a more general exploration of what it takes to have any variety of moral sensibility at all.

My suggestions have no doubt raised more questions than they have answered, but I hope one of the virtues to be found in my chapter is at least a satisfying response to the question of what we gain philosophically by studying atypical moral psychologies. The answer is quite simple. Real cases—especially difficult real cases—often force increasing conceptual sophistication where no amount of thought experiments will do the same. Hence, we may not gain so much an answer to long-standing philosophical debates as a realization that more traditional accounts are misconceived in important ways. If we are lucky, we also begin to see the direction in which, conceptually speaking, we need to move in order to amend these accounts; and with these amendments we likewise gain a better understanding of the kinds of empirical questions we have yet to pose.

With this apologia in place, I turn now to more specific replies to the three commentaries on my chapter. Although I can't address all the issues raised, I will do my best to respond to at least a substantial few so far as these fall into the two categories already mentioned: (1) those concerned with more general philosophical questions (chiefly from Kennett and Maibom), which are discussed first; and (2) those concerned more specifically with autism (chiefly from de Vignemont and Frith), discussed in the second section.

What Makes Us Moral Agents: Philosophical Considerations

I am particularly grateful to Kennett for succinctly stating and thus emphasizing the bottom-line philosophical position toward which the arguments of my chapter have tended. To wit: "[T]he terms of the debate between rationalists and sentimentalists must be modified. Recent evidence on moral development from the social and cognitive sciences and from psychopathy does not endorse the philosophers' traditional distinction between the affective and the cognitive, or their attempts to locate morality wholly in one or other domain" (this volume, p. 259). Indeed, let me say again in my own voice that what matters for our being moral agents—that is, for being the sort of moral agents we are—is that we are reasoning creatures with a certain range of affectively determined concerns. Take away either the affective component or the reasoning component and you take away our capacity for moral agency. However, this statement, which is true in its way, is also somewhat misleading. It continues to suggest that these components are related in such a way that it might be possible to subtract one or the other of them; that it might be possible to find one "spared" and the other "impaired," say, in autism or psychopathy. This certainly has been one popular way of characterizing these disorders. Empirical investigation, however, shows that each of these disorders involves impairments of both reasoning and affect, albeit impairments of different sorts. This suggests that we should shift our theoretical focus away from making too much of the divide between reasoning and affect and toward understanding why particular impairments of reasoning are bound up with particular impairments of affect—and, beyond that, why particular cognitive-affective profiles seem particularly detrimental to moral agency (e.g., as in psychopathy), whereas others seem to be less so (e.g., as in autism).

That said, I see nothing wrong with the conceptual project of trying to clarify analytically what each of these components contributes to the making of a moral sensibility. Here again I am grateful to Kennett for stressing the difference between (1) the empirical project of understanding how human beings are psychologically structured to be aware of and responsive to morally charged situations (Kennett suggests this was Hume's primary concern) and (2) the conceptual-normative project of understanding what it means for an agent to take a situation as morally charged, as generating normative reasons for action, reasons of the form "I ought (morally) to ϕ" (Kennett suggests this was Kant's primary concern). Of course these projects are not unrelated, since the former must surely act as a kind of negative constraint on the latter. Whatever we think is conceptually necessary, psychologically speaking, for taking situations as morally charged had better be instantiated by those human beings we count as moral agents. However, it may be that those human beings we count as moral agents are only a subset of the possible psychological types. So one consequence of pursuing the conceptual-normative project is that we gain a better understanding of the range of moral-psychological possibilities, and with that a better understanding of the kinds of individuals we should count as genuine moral agents.

What, then, are the psychological requirements for seeing situations as morally charged? Here again I am sympathetic to Kennett's claim that it

is not enough to have immediate desires or feelings (as animals might) that simply push and pull us about. That sort of psychology would make situations seem attractive or unattractive. Yet it would not allow for the kind of reflective and behavioral regulation that makes possible either (1) layering "I ought to ϕ" judgments (or cognates) over more immediate "I want to Ω" judgments (or cognates), which is required for moral evaluation; or (2) having "I ought to ϕ" judgments triumph over "I want to φ" judgments in producing action, which is required for moral behavior. By contrast, the kind of psychology that would allow for such feats of reflective regulation is one that according to Kennett incorporates a Kantian reverence for reason, which she takes to mean a regulating or limiting disposition to "seek and respond to normative considerations" (this volume, p. 260) or, alternatively, to "act in accordance with our reasons as we understand them" (this volume, p. 259). This kind of psychology would take situations in a normatively thick way, as generating reasons—what we might call regulative second thoughts—for thinking and acting in one way, even though our immediate desires, impulses, or feelings may sometimes pull us in a different direction altogether.

So far, so good. Now what precisely is involved in taking situations as generating reasons (regulative second thoughts) for us to act one way or another? One obvious point is that we must have certain look-ahead capacities, we must be able to calculate the consequences of doing (or not doing) different things. Yet that obviously is not enough. Even if we excelled at mapping out sets of consequences, we must be invested in certain particular outcomes for these calculations to eventuate in reasons to do (or not to do) the various things we contemplate. In a word, we must have future-directed ends—ends to which we are committed, ends that have the psychological power, therefore, of dictating what we ought to be doing, even sometimes against some current contrary impulses. Of course it stands to reason (hence, to the reason of reasoning agents) that the more coherent our ends, the stronger our reasons will be for or against doing any particular thing in the present. This is because the strength of our reasons will partly depend on their not speaking against one another—on their pulling as one in the same (or compatible) direction(s). Thus, so far as we are reasoning agents dispositionally structured to "seek and respond to normative considerations," it would not be surprising to find in us, in addition to the particular ends in which we invest, an interest in, or even a drive toward, making those ends as coherent as possible.

We now come to the nub of the issue: how to explain the fact of human beings coming to invest in particular future-directed ends, and especially

in those ends that are relevant to moral agency. Kantians have traditionally emphasized reason as a critical component; Humeans (and other sentimentalists) have traditionally emphasized affect. In the spirit of rapprochement, I agree with Kennett that both are necessary. Investment in particular ends, whether short or long term, whether involving the self or involving others, is for us an affective phenomenon, and the degree of our investment indicates the strength of our feeling, our care, for those particular ends. However, I certainly agree that such feelings are not just the crude affective buzzes we may sometimes get in our moment-by-moment interactions with the world. Rather, they constitute a new level of feeling, shaped and reshaped by reflection, in light of experience and anticipation, and continually subject to the pressures of becoming part of a coherent profile. We could call such feelings "reflective feelings" in order to acknowledge the shaping role of reason. However, this is not to suggest that, phenomenologically speaking, such "reflective" feelings need be experienced as any less "hot," any less immediate, any less strong than their more basic counterparts. If anything, given their etiology, such feelings will have more staying power; they are not mere whims of the moment. More important, they will have a regulative authority that stems from the way they survive in us reflectively, as part of the process of reasoning about the ends toward which we are affectively drawn.

Our next question is, what are the ends relevant to moral agency? What sort of cares and concerns must we have in order to regulate our short- and long-term behavior according to "oughts" that have a recognizably moral flavor?

Obviously, as Maibom insists, concern or compassion for others must be of central importance. This striking feature of human psychology has been well researched under the omnibus rubric of "empathy." I think this term is unfortunate since there are a variety of cognitive-affective phenomena ambiguously designated by it. For instance, what is sometimes called "empathy" is not care or compassion for others at all, but rather perspective-taking skills, which, to my way of thinking, can support and enhance our concern for others but are not fundamental to the existence of such a concern. Research on psychopathy and autism has been particularly useful in emphasizing the need for some disentanglement, since both disorders have been characterized as involving impairments of empathy, although obviously these impairments are of very different types. In fact, I think theorists might be well advised to abandon the notion of empathy altogether as a well-defined (or definable) construct in cognitive research. Failing that, we need to exercise considerable caution in treating it as a

unitary phenomenon usefully characterized as "spared" or "impaired." In any case, I have tried in my chapter to replace the notion of empathy with terminology that is no doubt still too crude, but which aims to be more precise in targeting the variety of concerns relevant to moral agency. So let me return now to a list of those concerns.

I begin, as I said, with care or compassion for others. I have speculated that this concern has its source in a distinct cognitive-affective system that develops naturally out of mechanisms responsible for early attachment and for the early recognition and attunement of emotions, but I agree that it is significantly enhanced by our more advanced perspective-taking skills. However, apart from this concern or compassion for particular others, I think there is another kind of concern for others that can sometimes look rather similar, namely, the concern that they be treated with the respect they deserve, given their place—or what ought to be their place—in the social order. Following Jonathan Haidt and other like-minded psychologists, I embrace the observation that many of our moral emotional responses are provoked by seeing individuals (including ourselves) undermining or supporting what we take to be the appropriate social order (guilt, shame, outrage, indignation, resentment, embarrassment, pride, complacency, and so on). Such emotional responses count as moral in this way of thinking because, as Haidt puts it, they have disinterested elicitors and disinterested action tendencies. In these cases, we react as we do, not because of our care or concern for particular others per se, but rather because we care about how individuals operate as social beings in a well-defined social structure. In a word, we care about the social structure in and of itself. We care that it is supported and maintained, and we are willing to punish and accept punishment when that social order is endangered or undermined. This strikes me as a different kind of concern from our care or compassion for particular others, and I speculate that it originates in quite a distinct affective-cognitive system, with its own particular phylogenetic and ontogenetic developmental history. Furthermore, although this concern is also supported and dramatically enhanced by our perspective-taking skills, I don't think such skills account for its existence any more than they do for the existence of our concern for particular others.

Does this exhaust the range of concerns that motivate specifically moral judgments and behavior? I have suggested not. Once we accept the idea that moral emotions should be functionally defined as emotional responses that have disinterested elicitors and disinterested action tendencies, then it seems clear that there is a range of such responses that manifest a

concern for something even beyond the social order—a concern with maintaining something like what I have called cosmic structure and position. This is the concern that my commentators (both official and unofficial) have found most puzzling. This is not surprising because it's the one most underspecified in my chapter and so most in need of further elaboration and defense. This is a future project, but let me just mention a few considerations that favor the idea.

Many moral codes, perhaps more prominently in ancient and nonwestern cultures, have a number of prohibitions or exhortations about how to live in harmony with a universal order. The concept of such an order is of something impersonal and transcendent, a lawful way of being that governs the whole of the cosmos, including the workings of the natural world and all of the entities (gods, humans, or otherwise) that might exist within it. A nice example comes from the writings of Pythagoras: "*Themis* in the world of Zeus, and *Dike* in the world below, hold the same place and rank as *Nomos* in the cities of men; so that he who does not justly perform his appointed duty may appear as a violator of the whole order of the universe" (Cornford, 1957, p. 12). This idea of there being a morally relevant order in the universe is not unique to ancient Greece. It is also contained, for instance, in the ancient Egyptian concept of *Maat*, the Persian concept of *Asha*, the Chinese concept of the *Tao*, the Vedic Indian concept of *Rita*, as well as the Hindu concept of *Dharma*. In all of these traditions there is a moral imperative laid upon human beings to understand and follow the precepts of the universal way as these pertain to the peculiarities of human existence. Thus, there are specific prescriptions about how to organize one's daily routines and rituals, including how and what to eat, what to wear, how to bathe, how to treat others, and so on and so forth—all supposedly derived from a proper understanding of this universal order. Hence, I disagree with both Maibom and Kennett that the human preoccupation with cosmic structure and position is not especially conducive to adopting precepts with specifically moral content. If anything, many actions that are taken to fall outside the moral domain in some cultures (e.g., our own) are moralized by others precisely because of the way they prioritize this sort of concern (e.g., cleansing rituals, vegetarianism, or treating the environment in a certain way).

How, more specifically, should we characterize the affective-cognitive system in which our concern with cosmic structure and position is rooted? Why should we have such a concern in the first place? In my chapter I suggested that it stems from the need to locate ourselves in a spatiotemporal order of things. In my conception, this need parallels the need to

locate ourselves in the social order of things and grows out of our uniquely human capacity to see the world (including ourselves) as extended in time. Now, it is interesting that even though Kennett doubts that the capacity for intertemporal perception is "specifically concerned with what most of us would take to be the content of human morality" (this volume, p. 261), she suggests that it may actually be fundamental to moral agency, since this is the capacity that allows us to conceptualize ends in the first place—ends in which we become affectively invested. I like this suggestion, but still I'm inclined to push it a bit further in order to explain why this intertemporal capacity can lead to a substantively moral worldview. In my conception, while this capacity gives us certain abilities, it also creates in us a particular need; namely, the need to make sense of ourselves in the larger scheme of things and hence to "discover" (i.e., impose) a cosmic order on things, just as we "discover" (i.e., impose) a social order on our immediate interpersonal environment. Moreover, as in the social case, the "discovery" (i.e., imposition) of cosmic order will encourage the formation of many rituals and routines geared toward supporting and maintaining that order. Thus, we see throughout human history the birth of many substantive cosmic moral orders.

Now here's an interesting possibility: If this account is on the right track, then it may help clear up a phenomenon that is otherwise quite mysterious. Prima facie, our concern with particular others and our concern with social order have the most immediate moral content; but then doesn't it seem odd that such concerns are frequently and blatantly sacrificed for the sake of some greater good? What greater good could there be? My answer is: maintaining the cosmic order. Following Kennett's suggestion, I propose that the reason we are so committed to serving such an end has much to do with the fact that the affective-cognitive system in which this concern is rooted is fundamental, in evolutionary and developmental terms, to our very existence as moral agents. As a result, it has a kind of priority that cannot be easily overruled.

Before leaving the topic of the variety of affective concerns that I say go into the makings of a (typically human) moral sensibility, let me clear up one important source of confusion. Maibom, in her comments, worries that my proposal is too inclusive in the following sense: Many thoughts and actions that would count as morally motivated in my view are in fact deeply immoral by most intuitive measures. For instance, citing the evidence of the Milgram and Stanford prison experiments, she points out that our powerful drive to conform to social roles, perhaps out of an abiding

concern for the social order, can lead us into "countless transgressions of ordinary human decency and competing moral norms" (this volume, p. 270).[1] Thus, Maibom questions whether it is really appropriate to count, for instance, our concern with social order as a genuine source of morality.

My response is that there seems to be an elision here between two different projects, and I'm grateful to Maibom for giving me the opportunity to disentangle the two. One project, which I take to be my own, is to explore what it takes to be any kind of moral agent at all, whether good or bad, i.e., the kind of agent that is an appropriate target for moral praise or blame. Such an agent, I claim, is one who must have certain capacities, both ratiocinative and affective, in order to be regulable by considerations that trump immediate and narrow self-interests. Such an agent must be capable of reasoning about ends toward which her activities tend, and she must be affectively invested in ends that make something other than her own well-being the focus of concern. A second project, reflected in Maibom's objection, is to consider what it takes for an agent to arrive at objectively correct moral judgments. What are the concerns an agent ought to have, or what should be the order among these concerns, for that agent to think and act in morally justified or praiseworthy ways? The latter project is concerned with delivering a substantive moral theory, whereas the former project is merely concerned with identifying the sorts of agents to whom such a moral theory could be appropriately addressed.

Now one might argue that unless an agent is moved to think and operate in accord with the correct substantive moral theory, she shouldn't count as a moral agent at all. This extreme view, which simply collapses the distinction between the immoral and the amoral, seems no more justified than an analogous view in the case of reasoning that would collapse the distinction between reasoning badly and not being in the game of reasoning at all. Of course, in the reasoning case, we can easily see that reasoning badly is an important phenomenon to investigate, especially from the perspective of understanding how the capacity for reasoning exposes less-than-ideal reasoners to certain kinds of liabilities that are entirely lacking in nonreasoning creatures. As George Eliot (echoing Hobbes) compellingly reminds us, it is "the power of generalising that gives men so much the superiority in mistake over the dumb animals" (Eliot, 1874/1996, p. 556). Likewise in the moral case, as I emphasize in my chapter, it is important to understand why our specifically moral interests and motives often drive us to acts of cruelty and destructiveness that have no place among the dumb animals.

Autism and Moral Agency: Conceptual and Empirical Considerations

Although my chapter in this volume was sparked by considering the problem of autistic moral agency, I will not say much on that topic here. In part this is because I agree so strongly with some of the points my commentators make, and in part because responding to certain other points would take me too far afield. For instance, in regard to the latter, de Vignemont and Frith make some fascinating remarks introducing a distinction between allocentric and egocentric representations of an agent's relationships with others, suggesting that both are involved in normal moral agency, but that the interaction between them has been "broken" in Asperger's syndrome. Consequently, individuals with this syndrome may display extreme egocentrism in their dealings with others, or extreme allocentrism (which I guess explains an apparently inflexible and disinterested commitment to rules, no matter what the consequences), but no shades of gray in between. That is to say, autistic individuals show little sign of motivating and modulating their rule following with the kind of egocentrically represented other-caring feelings that can be generated in particular situations because of how the plight of other affects the autistic person. In de Vignemont and Frith's account, it seems to be this sort of motivation and modulation that is necessary for genuinely moral behavior. As I said, I find this an interesting suggestion, but hesitate to comment in depth about how this proposal connects with my own without seeing a more detailed version. More mundane, I think, are the points on which I agree with de Vignemont and Frith, and I begin with these if only to bring the differences between our views into sharper focus.

As a way of introducing their own proposals, de Vignemont and Frith make the following summary claim: "it is misleading to characterize ASD as a lack of empathy associated with a preserved sense of morality" and, thus, explaining "the limitations in social and moral cognition in ASD patients require[s] a more subtle conceptual framework" (this volume, p. 277). Since this is precisely how I would summarize my own position, I take us to be engaged in similar kinds of exploratory conceptual projects, driven by the realization that received ways of characterizing autistic abnormalities are inadequate to what researchers are beginning to discover. Of course, we may have different views about the nature of the sophistications required: de Vignemont and Frith seem to favor a more purely cognitive approach to advancing our understanding of autistic motivation and behavior (i.e., by appealing to the need for a distinction between different types of representation), whereas I have suggested a need to develop our views in both affective and cognitive dimensions (i.e., by broadening our understanding of the range of concerns relevant to moral life and by seeing how these may be differently affected through an unusual profile of cognitive assets and deficits). Still, we agree on both these aspects of their negative claim: that autism should not be characterized as involving an "impaired" capacity for empathy, and/or a "spared" capacity for moral agency. Nevertheless, are the reasons for our agreement the same?

I have already voiced my own objection to any continued and unqualified use of the omnibus and ambiguous notion of empathy, so here just let me reiterate that it's precisely the kind of data that de Vignemont and Frith cite in connection with autism (and also psychopathy) that forces theorists to develop more precise theoretical constructs adequate to the task of distinguishing among the kinds of abnormalities manifested by these different populations; e.g., cognitive impairments in perspective-taking skills (as found in autism) versus impairments in at least some aspects of base-level affective responsiveness (as found in psychopathy). Perhaps it would be acceptable to retain the notion of empathy as long as theorists distinguish carefully enough between what de Vignemont and Frith refer to as the "cognitive and affective components of empathetic behaviors." This seems to be the preferred strategy adopted so far in the literature (for a review, see Hansman-Wijnands & Hummelen, 2006). However, I don't favor it myself because I think it encourages a tendency to characterize each of these components now as straightforwardly "spared" or "impaired." My bet is that this will also prove to be an unhelpful oversimplification insofar as normal empathetic development depends on the normal development of perspective-taking skills, and vice versa. Consequently, in my own positive account I have preferred simply to acknowledge that autistic individuals are, at some basic level, responsive to others' emotions, and then try to use this fact to account in part for the regulative concerns discernible in their reflection and in their activities.

Now what about the claim that autistic individuals show a "preserved sense of morality"? Once again, I agree with de Vignemont and Frith that this claim is misleading, and for many of the reasons they cite. As I point out in my chapter, autism is a spectrum disorder with individuals varying widely in terms of abilities and disabilities, even without factoring in issues of comorbidity; and for the very disabled end of the spectrum, it seems clear that no question of moral agency sensibly arises. What about those individuals who are relatively high functioning, i.e., where their autism is not associated with widespread and generally debilitating cognitive impairments? Here, too, I have argued in agreement with de Vignemont and Frith

that we need to be careful in attributing this to a "spared" moral sensibility, since behaviors that appear to be characteristic of moral judgment and moral motivation may be underpinned by rather different kinds of cognitive and/or affective processes. (This indeed was my point in observing that merely "passing" contrived theory-of-mind tests is no indication of a "spared" theory-of-mind capacity, since very able individuals may use compensating cognitive strategies for "hacking out" a correct solution to these sorts of problems—strategies which, by the way, do not fare so well in more naturalistic settings.) Thus, I agree that even though we see in many autistic individuals a drive to discover and follow various sorts of rules operative in our society, it remains an open question as to whether this drive indicates any deep understanding of why we have such rules, especially in those cases where typically developing individuals would understand the rules to have a specifically moral character. For many autistic individuals, I anticipate the answer would be "no," but surely not for all, as indicated by the anecdotal evidence of autistic self-report.

Is the evidence sufficient for reaching this sort of conclusion? Of course I agree with de Vignemont and Frith that speculations are not the same as conclusions based on broad-ranging and systematic studies. There are certain things we cannot say without having a great deal more data; for example, we cannot say much in a general way about autistic rule following. Still, general conclusions are not the only ones worth making. If my arguments are persuasive about what constitutes a moral sensibility, then as long as the reported self-reflections of someone like Temple Grandin are indeed her own reflections (and not, for instance, ghost written by someone else), it seems we have all the evidence we need to conclude that at least some high-functioning individuals with autism have a variety of moral sensibility. Would I call this sense of morality "intact" or "preserved"? Once again, my preference is not to use terms like these simply because, to my ear anyway, they imply something like normal functioning, and, as far as we can judge from the anecdotal evidence, autistic moral sensibility (where it exists at all) is quite unlike the moral sensibility found in typically developing individuals.

In sum, my views are perhaps not so distant from de Vignemont and Frith's as their commentary suggests. However, there are some critical points on which we do substantively disagree, and I would like to conclude by mentioning three of these. The first two, which involve only quick observations, bear on de Vignemont and Frith's conception of what constitutes a moral sensibility. The third point requires somewhat fuller elaboration because it involves their interpretation of certain data. All in all,

however, these remarks tend in the same direction, namely, toward more optimism than de Vignemont and Frith yet evince about the possibility of autistic moral agency.

What does it mean to be possessed of a moral sensibility, according to de Vignemont and Frith? One thing they explicitly mention is the ability to recognize, and of course respond to, violations of moral, as distinct from conventional, norms. But what are moral violations? They identify such violations with acts that lead to others' suffering, but immediately qualify that equation by saying that the suffering so caused must not be morally justified. My first point of criticism is that this account of moral violations is circular. It doesn't tell us how to recognize moral violations unless we already have a sense of what it is for certain acts to be morally justified. The second criticism is related. If some acts that cause suffering in others are morally justified, then this means that there are concerns other than concerns about others' suffering that are morally relevant, concerns by reference to which these acts are presumably justified. Thus, de Vignemont and Frith owe us a fuller account, even on their own terms, of the range of concerns properly involved in the manifestation of a genuine moral sensibility. Once these have been articulated, de Vignemont and Frith may actually find that the normative preoccupations observed in autistic individuals are to some degree manifestations of such concerns, arguing in favor of these individuals possessing a genuine variety of moral sensibility despite their somewhat attenuated understanding of others' suffering.

The third criticism that I want to make bears on de Vignemont and Frith's interpretation of some data relevant to the question of why autistic individuals comply with certain norms. As background to this point, let me be clear that we all agree that some moral offenses are offenses because they cause (morally unjustified) suffering in others, and equally we all agree that autistic individuals can be quite reliable in complying with norms that prohibit such harms. However, de Vignemont and Frith suggest that when autistic individuals comply with those norms, they most likely comply for the wrong reasons: not because of a true sense of suffering that those affected undergo, but rather because of a sort of unreflective norm worship—rule following for rule following's sake. As I have said, I don't rule out this possibility, but I worry that de Vignemont and Frith may be embracing this conclusion a little too quickly given their interpretation of some recent studies.

Consider their reaction to Blair's 1996 study indicating that autistic individuals (unlike psychopaths) are able to make the moral-conventional distinction much like normal controls, thereby seeming to demonstrate an

understanding of the moral import of certain norms (R.J.R. Blair, 1996). De Vignemont and Frith worry that because the subjects were only asked about the permissibility, the seriousness, and the authority jurisdiction of norm violations, the study is limited in what it can show. Specifically, it fails to rule out the possibility that autistic individuals have simply cottoned on to the fact that some transgressions are worse than others without truly understanding why. Other studies support such a possibility. For instance, Shamay-Tsoory and colleagues have shown that autistic individuals can be quite good at detecting when someone makes a faux pas, but they evince no understanding of why a faux pas is bad; i.e., according to de Vignemont and Frith, they evince no understanding that faux pas cause others distress (Shamay-Tsoory, Tomer, Yaniv, & Aharon-Peretz, 2002).[2] This concern may be further supported by reference to Temple Grandin's own case, where, in good anthropological style, she explicitly notes that some norm violations (which all involve social taboos) are treated more seriously than others, and despite the fact that she fails to understand the "logic" behind these prohibitions, she is committed to avoiding such "sins of the system."

Now I agree that these sorts of examples provide evidence for claiming that autistic individuals have (1) an interest in detecting and following different kinds of rules no matter what the rationale, and (2) an incapacity to understand why (typical) human beings should care about making or following at least some of these rules. However, I don't see that they support the stronger claim that autistic individuals are not capable of understanding the moral significance of some norms as far as this relates to harming others. Certainly Grandin herself is sensitive to the special quality of certain norms—for instance, against stealing, destroying property, and injuring other people. Even from a very young age she put these into a separate category from her so-called sins of the system.

Likewise, the faux pas study does not really support the idea of global autistic insensitivity to the wrongness of harming others. After all, even though others may suffer as a consequence of faux pas, there are really two counts on which one would not expect any deep understanding of this on the part of autistic individuals. The first is that faux pas usually involve norm violations having to do with respect for privacy, for social standing, or for some other aspect of social life to which autistic individuals are quite oblivious (cf. Grandin's failure to understand the rationale for certain social taboos, her "sins of the system"). The second count is that insofar as faux pas cause suffering, the sort of suffering in question is usually more psychological than straightforwardly physical, consisting in

a range of highly developed social emotions—guilt, shame, embarrassment, and the like—that autistic individuals have little experience of themselves and difficulty detecting in others. Thus, autistic individuals may well be insensitive to the specific phenomena of harms caused in certain situations, but this, to echo Jeanette Kennett's earlier claims, says more about their incompetence as moral agents than it does about their being out of the game of moral reflection and regulation altogether.

Notes

1. Maibom also raises some interesting questions about the morally questionable phenomenon of obedience. I agree with her that the evidence shows that typically human beings are psychologically geared to defer to authority and that such deference is a mixed blessing. Clearly it makes us capable of living together in social groups, conforming to expectations without the need for a lot of heavy-duty threats or other mechanisms of compliance. However, such deference also has a downside, as the Milgram and Stanford prison experiments make clear. Yet how is all this connected with moral agency? I think there is no simple answer to this question because following rules, or deferring to authority, can clearly be done in different ways (as I indicated in my discussion of autistic rule following). For instance, one might defer to a rule or to some authority "mindlessly," as we might say; that is, one just automatically defers, no matter what (some autistic rule following may fall into this category). In my view, this is not the stuff of morally agential behavior, and perhaps what these experiments show is that even typically developed human beings are all too ready to abjure any semblance of such behavior.

However, when it comes to obedience, there are also other possibilities. Perhaps one defers on a particular occasion because one thinks it's the "right" thing to do. Now we are getting into the area of morally agential behavior, but this too comes in degrees. For instance, one might think something is the right thing to do because someone in authority said so, and the right thing to do is to defer unquestioningly to authority. This is clearly less agential than thinking one ought to defer to authority on some occasion because that authority has better access to determining what is independently the right thing to do. Alternatively, also more agentially, one might defer to authority on some occasion because that authority happens to dictate what one independently thinks is the right thing to do. Or one might defer to authority on some occasion because one independently values the sort of social structure in which authority is paid a certain amount of deference (within limits). I mention all these possibilities simply to emphasize the point that the relationship between obedience and moral agency is not straightforward, raising a host of interesting issues that are somewhat orthogonal to the main themes of my chapter and deserve far greater attention than I can give them here.

2. Are faux pas appropriately seen as wrong because they cause distress? I myself am sceptical of this analysis. I suspect that what makes such acts wrong is that they transgress socially accepted norms, and that whatever pain they cause is not primarily a function of the acts in and of themselves but instead a function of the fact that these acts are seen to be socially transgressive. In other words, faux pas are instrumentally rather than constitutively distressing insofar as their distress-causing properties are contingent upon an individual's understanding and acceptance of the social norms they transgress.