

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 Varieties of Moral Agency: Lessons from Autism (and Psychopathy)

Victoria McGeer

The Roots of Moral Agency

What makes us moral creatures? What are our moral concerns? How do we come to make moral judgments and act from specifically moral motivations? These fundamental yet perplexing questions, once solely the province of philosophers and theologians, have been pursued with increasing interest by psychologists, anthropologists, evolutionary theorists, ethologists, and lately, cognitive neuroscientists, yielding new and often challenging insights into this critical aspect of our human condition.

We are unquestionably unique in the shape and quality of our moral experience and behavior—the “paragon of animals,” as Shakespeare has Hamlet say, “in action how like an angel in apprehension how like a god!” And yet because of these qualities—i.e., specifically moral experiences, projects, and concerns—we are likewise capable of immense and ingenious evil. Here, then, is a first and familiar paradox of our moral existence: We could not be as ungenerous in our condemnations, as cruel in our fanaticisms, as self- or other-destructive in our pursuits, if our lives were not animated through and through with moral purpose and concern. Moral agency is a double-edged sword and seems inevitably so, no matter how much we trumpet the glories of our moral sensibility. Why should this be the case? Is it because our moral capacities are fundamentally rooted in our sentimental natures—as Hume or Smith would maintain—so that we depend on various affective states to move us to care appropriately for, and sometimes inappropriately about, one another? Or are our moral capacities fundamentally reason-based or “cognitive,” as Lawrence Kohlberg would say, where reason is given the dual job of arriving at appropriate moral conclusions and of channelling our “affective forces” to ensure that we act in accord with our moral judgments (1971)?¹ In this case, iniquitous behavior might stem from a failure of reason either to arrive at

appropriate moral judgments or to control our affect in a morally acceptable way.

The relation between reason, emotion, moral judgment, and behavior is an old and contested one, with paradigm and opposing positions associated with Hume on one side and Kant on the other. Central to this theoretical division is the role of what is now often called "empathy"² in generating the kind of concern for others that motivates and regulates paradigmatic instances of our moral behavior. Hume, in keeping with his view that "the ultimate ends of human actions can never . . . be accounted for by reason, but recommend themselves entirely to the sentiments and affections of mankind" (Hume, 1777/1975, p. 293), argued that the capacity to feel with and like another—to enter sympathetically into their cares and concerns—was critical for developing and maintaining an other-regarding moral agency. Kant, by contrast, was deeply disdainful of the moral importance of empathy and/or sympathy, favoring a moral psychology motivated in its purest form by a rational concern for doing one's duty:

If nature had implanted little sympathy in this or that man's heart; if (being in other respects an honest fellow) he would be cold in temperament and indifferent to the sufferings of others—if such a man (who in truth would not be the worst product of nature) were not exactly fashioned by her to be a philanthropist, would he not still find in himself a source from which he might draw a worth far higher than any a good-natured temperament can have? . . . For love out of inclination cannot be commanded; but kindness done from duty—although no inclination impels us—is *practical* not *pathological* love, residing in the will and not in the propensities of feeling, in principles of action and not of melting compassion; and it is this practical love alone which can be the object of command. (Kant, 1785/1948: 398–389, as cited in Kennett, 2002, pp. 352–353; emphasis in original)

One fruitful way to pursue this debate in a contemporary context is by looking at the moral capacities of individuals in various clinical populations. There we may expect to discern how particular abnormalities in cognitive and/or affective capacities compromise moral agency. Two populations of particular interest in this regard are autistic individuals and psychopathic individuals. Both these populations seem to lack empathy in some sense of that word, and yet psychopathic individuals are well known for their lack of moral concern whereas individuals with autism can have strongly felt moral convictions despite the fact that their moral judgments are often impaired by the difficulties they have in understanding other points of view. What accounts for this difference?

In trying to resolve this puzzle, I begin by reviewing how the debate on empathy has unfolded in light of lessons that theorists have taken from the study of psychopaths. I then turn to the problem of autistic moral

concern, basing my discussion on an insightful and challenging paper by Jeanette Kennett. Here I will agree with Kennett that even though a focus on empathy has yielded important insights into our moral nature, such an approach is also restricted in how much it can explain. However, pace Kennett, I resist the Kantian conclusion that "reverence for reason is the core moral motive" (Kennett, 2002, 355). Rather, I will conclude that the concern with affect has been too narrowly focused on empathy, and that this has stemmed in part from a persistent tendency in philosophy and other academic inquiries to try to locate the essence of our moral nature in a single cognitive capacity or affective disposition. This tendency is not as strong as it once was, but nevertheless it continues to shape contemporary debate.

Still, while old habits die hard, the evidence from development, psychopathology, cross-cultural studies, and even primate studies has been pushing us in a different direction for some time (Haidt, 2001; Shweder & Haidt, 1993; Shweder, Much, Mahapatra, & Park, 1997). And I think the evidence from autism supports this trend, for here I think we find some persuasive indication that our moral nature is shaped by (at least) three different tributaries of affectively laden concern which I tentatively label as follows: (1) concern or compassion for others, growing out of the attachment system and fostered mainly by a capacity for emotional attunement between self and other, although later also supported by perspective-taking skills; (2) concern with social position and social structure, growing out of the need to operate within a hierarchically organized communal world and fostered by our highly developed perspective-taking skills; and finally, (3) concern with "cosmic" structure and position, growing out of the need to bring order and meaning to our lives and fostered by our capacity to view ourselves in intertemporal terms. Although I won't be able to elaborate on these three spheres of affective concern in any great detail, my view is that they are responsible for producing the cross-cutting systems of value that give shape to our moral being, a shape that can be differentially influenced by affective and/or cognitive impairments that target any or all of these spheres of concern. This is what we see in both individuals with autism and individuals with psychopathy, even though the nature of their impairments gives rise to very different moral psychological outcomes.

The Importance of Empathy: Lessons from Psychopathy

Psychopaths have long been known for their apparent amorality, specifically for their deep indifference to the cares and sufferings of others, leading them to act in cruel and often criminal ways.³ Theorists have posed

the following question: Is this primarily an emotional deficit that prevents psychopaths from empathizing or affectively experiencing the world from other points of view and so from taking account of how their activities may negatively affect others? Or is it some kind of cognitive deficit that makes psychopaths unable to see why the cares or concerns of others should matter to them, even though they are perfectly well aware of these cares and concerns?⁴

One noteworthy feature of psychopathic individuals is their apparent facility with mental state attribution. In sharp contrast to individuals with autism, for instance, they seem remarkably adept at reading the minds of others, if only to manipulate them. They are glib and frequent liars, passing themselves off with an easy charm that speaks to a ready, though perhaps superficial, understanding of social norms and expectations (Cleckley, 1955; Hare, 1993). This clinical impression is borne out by their normal performance on standard and advanced theory of mind tests (R.J.R. Blair, Sellars, Strickland, Clark, Williams, Smith, & Jones, 1996). Consequently, they appear to have no cognitive deficit in understanding others' states of mind, including their beliefs and desires, motives and intentions, cares and concerns. In one straightforward sense of this term, they have no difficulty with perspective taking (Nichols, 2002a).

In contrast with this cognitive capacity, psychopathic individuals have been found to be notably abnormal in their affective profile. Clinical reports indicate that they show "a general poverty of major affective reactions, particularly those that would be triggered by the suffering of others (remorse, sympathy), condemnation by others (shame, embarrassment), or attachment to others (love, grief)" (Haidt, 2001, p. 824; cf. Cleckley, 1955; Elliott, 1992; Hare, 1993). They have trouble recognizing some facial expressions of emotion, especially fear and sadness (R.J.R. Blair, Colledge, Murray, & Mitchell, 2001b; Stevens, Charman, & R.J.R. Blair, 2001). They also show an abnormal autonomic reaction to these emotions, responding to the distress cues of others (facial and vocal expressions) as if they were affectively neutral (R.J.R. Blair, Jones, Clark, & Smith, 1997). What impact might this lack of affective responsiveness, especially to distress, have on their moral capacities?

James Blair has argued that it likely has a significant impact, explaining, for instance, why psychopaths fail to distinguish between moral and conventional transgressions (R.J.R. Blair, 1995). Making such a distinction is now viewed as a critical indicator of moral capacity—of being able to regard activities in a specifically moral light and to make judgments about those activities that have a specifically moral character (Nucci, 2001;

Smetana, 1993; Turiel, 1979, 1983; Turiel, Killen, & Helwig, 1987; but for criticism, see Kelly, Stich, Haley, Eng, & Fessler, 2007). It is a distinction that is made cross-culturally (Hollis, Leis, & Turiel, 1986; Nucci, Turiel, & Encarnacion-Gawrych, 1983; Song, Smetana, & Kim, 1987) and begins to emerge in normally developing children from around the age of 39 months (Smetana, 1981). In normal populations, moral transgressions are characteristically regarded as rule or authority independent, i.e., they are viewed as wrong whether or not there is a rule proscribing them, whether or not someone in authority licenses them. Conventional transgressions, by contrast, are normally regarded as wrong only because the acts are proscribed by rules of acceptable social behavior; if the rules were changed or suspended, the proscribed acts would no longer count as wrong (Turiel, 1983).⁵ One explanation for this is that moral transgressions are normally regarded as wrong—and more seriously wrong (less permissible) than other sorts of transgressions—because they provoke a strong affective response in us (Haidt, Koller, & Dias, 1993; Nichols, 2002b, 2004b). Thus, for instance, we code those transgressions that result in victims, individuals who suffer either physically or psychologically, as paradigm moral transgressions because of our affective response to the victims' imagined distress, something to which the psychopath is apparently blind. Hence, they fail to distinguish these sorts of moral transgressions from those that merely break the accepted rules of social life.⁶

In sum, this work on psychopaths seems to support the view that the capacity for moral thought and action is strongly dependent on our affective natures and in particular the capacity to respond empathetically to others' affective states, to experience a vicarious emotional response to how they affectively experience the world, and especially to feel some distress at their distress and suffering (see also Nichols, 2004b). This seems to speak to a Humean rather than a Kantian view of the roots of moral understanding and moral motivation. We develop a special concern for others and for their well-being because, as Hume says, "the minds of all men are similar in their feelings and operations. . . . As in strings equally wound up, the motion of one communicates itself to the rest; so all the affections readily pass from one person to another" (Hume, 1740/1978, pp. xix, 743). Or, in the words of contemporary psychologists Andrew Meltzoff and Keith Moore, "We 'do unto others' in a special way because there is a deeply felt equivalence between self and other. Without a sense of like-me-ness, we do not think our folk psychology and moral judgments would take the form that they do" (Meltzoff & Moore, 1999, p. 11).

The Limits of Empathy in Explaining Moral Agency: Lessons from Autism

Many people with autism are fans of the television show *Star Trek*. I have been a fan since the show started. When I was in college, it greatly influenced my thinking, as each episode of the original series had a moral point. The characters had a set of firm moral principles to follow, which came from the United Federation of Planets. I strongly identified with the logical Mr. Spock, since I completely related to his way of thinking.

I vividly remember one old episode because it portrayed a conflict between logic and emotion in a manner I could understand. A monster was attempting to smash the shuttle craft with rocks. A crew member had been killed. Logical Mr. Spock wanted to take off and escape before the monster wrecked the craft. The other crew members refused to leave until they had retrieved the body of the dead crew member. To Spock, it made no sense to rescue a dead body when the shuttle was being battered to pieces. But the feeling of attachment drove the others to retrieve the body so their fellow crew member could have a proper funeral. It may sound simplistic, but this episode helped me finally understand how I was different. I agreed with Spock, but I learned that emotions will often overpower logical decisions, even if these decisions prove hazardous. (Grandin, 1995, pp. 131–132)

This passage, written by Temple Grandin, a remarkably able individual with autism, articulates a puzzle for the view of moral agency that has been emerging thus far. For here and in other writings she combines her sensitivity and attraction to the existence of a moral order with her acknowledgment that she lacks the normal emotional profile of other human beings, specifically “the feeling of attachment” that drives others, for instance, to endanger themselves for the sake of a comrade, dead though he may be. Could it be that individuals with autism are lacking the basic kind of empathetic connection with others so far identified as being critical to the development of an other regarding moral concern? If so, why are they not like psychopaths in their callous disregard of others or, at the very least, in their insensitivity to the moral domain? Yet, on the contrary, as far as they are cognitively able, individuals with autism seem remarkably prone to view their own and others’ behavior in moral terms; i.e., in terms of duties or obligations that ought to be binding on all people, even if their sense of the nature of these duties and obligations can seem naïve or bizarre from our point of view.

Consider, for example, the case of a young man with perfect pitch and a passion for pianos who could not fathom how anyone could be happy without a well-tuned piano. Upon discovering that there were people who in fact didn’t have pianos, or who kept them out of tune, he thought there should be a constitutional amendment requiring every home to have a

well-tuned piano (M.A. Dewey, 1992, discussed in Kennett, 2002). His aim was clearly not to benefit himself, but rather to improve the lives of those around him and thereby make the world a better place. Of course, his moral priorities are dramatically affected by his autism. Nevertheless, from this and many other such examples, it seems clear that individuals with autism can and often do have a strongly developed moral sensibility or, as Kant would say, a concept of duty deriving from “consciousness of a moral law” (Kant, 1797/1991, p. 400). Where does this moral consciousness come from?

In her groundbreaking paper, Jeanette Kennett (2002) argues that the fact that individuals with autism are often deeply motivated to do the right thing should make us rethink how critical the capacity for empathy is to the development of moral agency. For, as she says, they seem in many ways even worse off than individuals with psychopathy regarding their ability to connect with other people. Psychopaths at least have relatively intact perspective-taking skills, readily surpassing even the most able autistic individuals in detecting others’ states of mind. And although psychopaths may have trouble seeing other people as fully real (Elliott, 1992, p. 210), perhaps because they lack any deep understanding of the range of emotions that animate them, the outsider status they experience in consequence seems less dramatic than what autistic individuals have reported about themselves, often saying that they feel like aliens beamed in from another planet, or, in the words of Temple Grandin, like “an anthropologist on Mars.”

Of course, there are large individual differences among people with autism. Characterized as a spectrum disorder, autism can be diagnosed in individuals who are relatively low functioning (their autism is combined with other mental handicaps) all the way through to those who are relatively high functioning (with normal to high IQ, often good, although characteristically abnormal language skills, and often compensating cognitive strategies for coping with their autistic disabilities). However, despite this wide range of ability, individuals with autism show a characteristic triad of impairments, according to which a diagnosis is made. This includes: (1) a qualitative impairment in reciprocal social interactions, including a marked lack of awareness of others’ feelings, abnormal comfort-seeking behavior at times of distress, impaired imitation, aversion to or abnormal physical contact, and lack of social play and peer friendships; (2) a qualitative impairment in nonverbal and verbal communication, including lack of eye gaze and facial expressions to initiate or modulate social interactions, abnormal prosody, echolalia, extreme literal-mindedness, and general difficulties with conversational pragmatics; and (3) impairments in

imaginative abilities, including lack of pretend (especially role-taking) play in childhood, highly restricted and repetitive interests, and an obsessive insistence on routine and environmental stability.

Although autism is not usually diagnosed until around 18 months, what makes it particularly noticeable is the relative absence of all those behaviors by which typically developing children normally register their sense of the vast difference between people—i.e., subjectively animated creatures “like me”—and other things in their environment. Such behaviors include joint attention; reciprocal imitation games; social referencing behavior, where toddlers use the affective expressions on others’ faces to guide their interactions with unknown objects, and so on (for a general discussion and review, see Baron-Cohen, Tager-Flusberg, & Cohen, 2000; Frith, 1989; Happé, 1994b). Indeed, this missing recognition of others “like me” can be so deep that one high-functioning adult would later remark: “I really didn’t know there were other people until I was seven years old. I then suddenly realized that there were other people. But not like you do. I still have to remind myself that there are other people. I could never have a friend. I really don’t know what to do with other people, really” (Hobson, 1992, p. 165, cited in Kennett, 2002).

What, then, is the source of autistic moral concern, since empathy in the sense of affective attunement with other people seems clearly beyond the scope of their experience? Kennett suggests that the answer can be found in focusing on autistic rationality, specifically autistic individuals’ susceptibility to and deep interest in the sense-making pull of reason. Indeed, she proposes that their “moral feelings are of a Kantian, rather than a Humean, cast” (2002, p. 352) since they seem to derive from a deeply felt practical concern to do the right thing, whatever that should turn out to be. We saw evidence of this Kantian sensibility in the Grandin passage quoted earlier, in which she explains that her liking for *Star Trek* is based on two things: first, that the characters in the show had “a set of firm moral principles to follow”; and, second, that she could identify with one of the characters in particular, the logical Mr. Spock, whose recommendations and behavior are guided by reason, not emotion. Grandin and other high-functioning individuals with autism seem particularly committed to the Kantian idea that their behavior (and everyone else’s) should conform to a principle of reason that includes them in its scope. Thus, they are prepared to see other people’s interests as reason giving in the same way as their own, even though, as Kennett says, their problems with perspective taking give them “great difficulty in discerning what those interests are” (2002, p. 354). Still, their apparent need to figure out the

“right” thing to do based on taking the concerns and interests of others into account leads them to make quite extraordinary efforts to understand those concerns and interests. For instance, Temple Grandin writes of having built up a “tremendous library of memories of experiences, TV, movies, and newspapers” that she consults in order to understand what others might be up to and so guide her social behavior appropriately (Grandin, 1995, p. 137). Another very able individual, Jim Sinclair, writes of his need to develop a “separate translation code for every person I meet” (Sinclair, 1992, p. 300).

From such examples we seem to have clear evidence that even though a lack of empathy makes it extremely challenging for individuals with autism to act in morally appropriate ways, it does nothing to undermine their interest in so acting; it does nothing to undermine their moral concern. Yet if we turn our attention to psychopathic individuals once again, it now seems puzzling that their apparently less dramatic lack of empathy should so gravely undermine their capacity for moral concern. This puzzlement can be dispersed, Kennett suggests, once we recognize that contemporary theorists have become overly focused on the affective dimension of moral life, thanks no doubt to the tremendous importance of empathy in guiding normal human relations. For just as the pull of reason can explain the autistic individual’s moralism, it seems likely that the psychopathic individual’s amorality might well be explained in terms of a failure of reason to operate in him with its normal motivational force (for a similar approach, see Malborn, 2005). Thus, Kennett proposes: “It is not the psychopath’s lack of empathy, which (on its own, at any rate) explains his moral indifference. It is more specifically his lack of concern, or more likely lack of capacity to understand what he is doing, to consider the reasons available to him and to act in accordance with them” (2002, p. 354).

This failure of reason may seem surprising. After all, our image of the psychopath is of a person who is rather good at serving his own interests without care or concern for the damage he does to others; hence of someone who is rather good at thinking and acting in instrumentally rational ways. However, as Kennett argues, this image is misleading, ignoring the dramatic ways in which psychopaths are also compromised in their ability to make sound prudential judgments. As Carl Elliott observes: “[W]hile the psychopath seems pathologically egocentric, he is nothing like an enlightened egoist. His life is frequently distinguished by failed opportunities, wasted chances and behavior which is astonishingly self-destructive. This poor judgment seems to stem not so much from the

psychopath's inadequate conception of how to reach his ends, but from an inadequate conception of what his ends are" (Elliott, 1992, p. 210, cited in Kennett, 2002). Thus, in Kennett's view, the psychopath is not able to regard others' interests as reason giving for him because he is not able to sustain a sense of why any interests should be reason giving, apart from the very short-term impulses that drive him from one action to the next. Perhaps it might be better to say that while the psychopath may have action-guiding impulses, he has no impulse-controlling interests since he has no "extended and coherent conception of his own or others' ends" from which such interests could be rationally derived (Kennett, 2002, p. 355). In this respect, high-functioning individuals with autism seem far better off. Their disabilities may make it difficult for them to form a clear or sophisticated conception of their own or others' ends, but they show a clear drive to give rational shape and meaning to their lives and to their interactions with other people. Hence, Kennett suggests, insofar as individuals with autism have "a basic conception of justification and of interests as reason-giving" (2002, p. 355) we can find in their psychological makeup a sufficient basis for moral agency, even if their lack of affective attunement leaves them rather unskilled in the moral domain.

Here, then, is the general conclusion to which Kennett thinks we are driven. On the one hand, if we focus on the moral limitations of individuals with autism, then it seems quite right to insist with the Humeans that affective attunement is hugely important for the development and operation of "autonomous, responsive, moral agency in human beings" (p. 357). Without any real sensitivity to the wide variety of pains, pleasures, and other emotions individuals experience under the myriad circumstances constituting normal social life, individuals with autism are seriously disadvantaged in developing a sophisticated understanding of the moral domain, either of the kinds of rules that ought to govern one's moral behavior (witness the young man with the passion for pianos) or indeed of the way rules ought to be applied, sometimes even set aside, to serve deeper moral ends. On the other hand, if we look at the moral capacities of individuals with autism, then it seems we must conclude that affective attunement is not necessary for the development of a genuine moral sensibility. Thus, Kennett suggests, "the story of how we normally get to be moral agents and the story of what is *required* for moral agency is not the same" (2002, p. 357). The social-cognitive abnormalities that distinguish autism from other developmental disabilities ensure a developmental trajectory that is not just delayed but deeply eccentric compared with a typical developmental trajectory. Nevertheless, at least among high-

functioning individuals with autism, their relatively intact reasoning abilities, coupled with a drive for order and a need to make sense of their own and others' behavior, seems to support the emergence of a sense of duty or conscience that is, by contrast, entirely lacking in the psychopathic population. Thus, a comparison between these two atypical groups of individuals suggests to Kennett that it is Kant, not Hume, who has put his finger on "the essence of moral agency, the concern to act in accordance with reason which animates agency and which we cannot do without" (2002, p. 355). To be empathetically insensitive to others is to be seriously disadvantaged in the moral domain, but to be insensitive to reason—that is, insensitive to reason as generated and sustained by various self-organizing ends—is to fall out of the moral domain altogether. Hence Kennett concludes in keeping with Kant that "reverence for reason is the core moral motive, the motive of duty" (p. 355).

Reply to Kennett: Does Autistic Moral Agency Show Us that Reverence for Reason Is the "Core" Moral Motive?

There is no doubt that Kennett raises an important issue for a broadly Humean account of the roots of moral agency. In particular, she makes clear to us that any account of moral agency must attend to what I will call the "agential" side of these capacities as much as to their moral side. That is to say, for anyone to be a moral agent, they must at least be a certain kind of rational agent—an agent who is capable of controlling their immediate impulses in the service of some larger end; hence an agent to whom reason speaks. However, in Kennett's view, such an agent is the kind of agent in which reason has its own motivational force: "[O]nly individuals who are capable of being moved directly by the thought that some consideration constitutes a reason for action can be conscientious moral agents" (2002, p. 357). This is a strong conclusion, perhaps stronger than it needs to be in order to preserve Kennett's critical insight about the agential side of moral agency. So, in this section, my aim is to review some further evidence from autism to bolster a certain aspect of Kennett's insight (that responsiveness to reason may be necessary for moral agency), while at the same time questioning the Kantian spin she puts on it (that responsiveness to reason is sufficient for moral agency). At the very least, I will argue that Hume's emphasis on various kinds of affect must not be abandoned too quickly. If we are to understand how certain ends become salient enough to compel our reason, thereby giving it some long-range appetitive control.

I begin with a word of caution. I have suggested that Kennett's conclusion may be overly strong in the modality of its claim about reason. There is another way in which it may be overly strong: namely, in its generality. Consider the form of Kennett's argument: (high-functioning) individuals with autism are unlike psychopaths in manifesting some degree of moral concern; therefore, despite the difficulties they have with moral judgment and moral behavior, individuals with autism must have some quality or capacity psychopaths lack that is "essential to the nature of [moral] agency" (2002, p. 357). The more general interpretation of this argument, intended by Kennett, is that a consideration of autistic moral capacities shows us something about the basic structure of moral agency *simpliciter*. Specifically, as she says, it shows us that a "reverence for reason is the core moral motive, the motive of duty" (p. 355). That is to say, anyone who is lacking in this reverence for reason will fail to be a moral agent. However, there is a less general interpretation of this argument: that merits attention: namely, that a consideration of autistic moral capacities shows us something about the special structure of autistic moral agency. Specifically, it shows us that a certain reverence for reason can go some way toward compensating for the lack of empathetic attunement that is essential for the development of a typically structured moral agency. That is to say, anyone who is lacking in empathetic attunement and who lacks this reverence for reason will fail to be a moral agent. The more general interpretation says that reverence for reason is the core moral motive for all individuals. The less general interpretation says it plays a particular kind of compensating role in individuals with autism.

Why favor the less general interpretation? Consider a close analogy. As we have already noted, autistic individuals are greatly handicapped in the social domain. One way of characterizing their primary deficit is in terms of an inability to represent others' and possibly even their own mental states. They lack what is often termed a natural theory of mind, i.e., a disposition, fine-tuned through development, for simply reading off from others' expressions and (contextually situated) behavior the mental states that motivate and direct them. One classic way in which this deficit is manifested is in the so-called false-belief task, where subjects are required to attribute a false belief to a character in a story in order to predict what she will do next. Normally developing children begin to pass this test by the time they are 4 years old and developmentally delayed children by the time they have reached a corresponding mental age. This is not true of children with autism. They continue to experience difficulty with this task even at a much greater mental age. For instance, from a large sample of

seventy autistic children compared with seventy normally developing children, Francesca Happé has shown that normally developing children have a 50 percent chance of passing false-belief tasks by the verbal mental age of 4, whereas autistic children have a 50 percent chance of passing only by the verbal mental age of 9.2 years (Happé, 1994b, pp. 71-73). Still, even though children with autism take more than twice as long to reach the same probability of success on this task as typically developing children, a "talented minority" will eventually pass, signaling that they have some capacity to represent others' mental states.

How can these results be explained? One possibility is that these talented children with autism have developed, after much delay, a relatively normal capacity for reading other minds. However, as many theorists have noted, the autistic capacity does not generalize easily to naturalistic settings. Furthermore, it is hard to say the capacity is still simply immature, since the kind of mistakes in mental state attribution that these individuals continue to make are very unlike the mistakes made by typically developing children at a much earlier age (Happé, 1994a). Hence, the more likely explanation of autistic "mindreading" is that some intellectually gifted individuals are able to use their advanced reasoning skills to "hack out" a solution to the puzzle of other minds, even while they continue to have no immediate or natural perception of others' mental states. As Temple Grandin says, "I have had to learn by trial and error what certain gestures and facial expressions mean" (Grandin, 1995, p. 135). In sum, the social cognition of individuals with autism remains extremely limited, with frequent and bizarre errors demonstrating a far from normal (albeit delayed) developmental trajectory. Thus, learning about the mechanisms of autistic social cognition tells us more about what the mechanisms of normal social cognition are *not* than about what they actually are. Perhaps the same will be true for the mechanisms underlying autistic moral concern.

Now let us look more closely at the abilities and disabilities of individuals with autism. What evidence might suggest that their moral sensibility is structured quite differently than the moral sensibility of typically developing individuals? To begin on a slightly downbeat note, recall that one feature of the autistic profile is an inflexibility of behavior that stems from an obsessive regard for rules and routines. Individuals with autism seem to have a great need to impose order on the world, no doubt because of neurological abnormalities that give rise to a disorienting, highly complicated, anxiety-inducing range of experiences. Clinical observations coupled with persistent subjective reports testify to an array of sensory abnormalities (auditory, tactile, olfactory, visual, nociceptive) in many individuals

with autism that make interacting with their environment extremely challenging. Indeed, they often find the mere physicality of other people's presence and/or social demands oppressive, such as being required to meet another's gaze or putting up with a "terrifying" embrace. On top of that, their difficulties in processing social and expressive cues make other people's behavior unpredictable and often overwhelming. Unsurprisingly, then, high-functioning individuals who can talk about what their autistic experience is like commonly report that fear and anxiety are their dominant emotions (Grandin, 1995, pp. 87–89). In eloquent testimony of this, Therese Jolliffe writes,

Normal people, finding themselves on a planet with alien creatures on it, would probably feel frightened, would not know how to fit in and would certainly have difficulty in understanding what the aliens were thinking, feeling and wanting, and how to respond correctly to these things. That's what autism is like. If anything were suddenly to change on this planet, a normal person would be worried about it if they did not understand what this change meant. That's what autistic people feel like when things change. Trying to keep everything the same reduces some of the terrible fear. (Jolliffe, Lansdown, & Robinson, 1992, p. 16)

Rules and routines help keep things the same, making the world emotionally and cognitively more approachable. To this end, autistic individuals are highly motivated to follow rules and are very concerned that others do so too. My downbeat suggestion, then, is that a good part of the behavior we identify as manifesting moral sensibility among individuals with autism may stem from a need to abide by whatever rules they have been taught without sharing our understanding of the ends those rules are meant to serve. In other words, for many such individuals, it may well be an open question as to how deeply their "moral" judgments and behavior are genuinely guided by moral concerns.

This issue is nicely illustrated by the following anecdote reported in one of the few studies of moral reasoning among autistic adults: "A young man with autism was participating in a board game called 'Scruples' which involves listening to stories and telling what you would do in each situation. He was given a scenario in which a store owner saw a woman stealing a small amount of food from his store. The store owner knew that this particular woman had no job, no one to support her, and several young children. The young man with autism was asked what he would do in the situation. He replied, 'Everyone has to go through the checkout line. It is illegal not to go through the checkout line. She should be arrested'" (Keel 1993, p. 49). Was this autistic man simply unable to comprehend the

woman's need-driven motivation and hence incapable of seeing that a milder response was called for? Keel, who reports the case, favors this interpretation. She writes,

This reply certainly seems cold and uncaring. However, it reflects this young man's social cognitive deficits in perceiving the intent behind another's actions. Certainly, he had never been in a situation where he could not afford his groceries. He goes to the grocery store weekly and always goes through the checkout line. Additionally, he has always been taught not to steal. Without the ability to appreciate the perspective of another or to consider intent as well as consequence, he appeared unable to give any other answer. (1993, p. 49)

It is surely true that this young man lacked sufficient perspective-taking skills to understand the woman's probable state of mind. Yet the interesting question remains: How deep is his understanding of the prohibition against stealing? Does it rise to the level of a genuinely moral understanding, where rules are followed not just because they are rules, but rather because they serve some deeper moral end? There may be reasons to doubt this. For instance, it seems relatively clear from the anecdote that, even without perspective-taking skills, the young man knew that the woman had children to feed and no money to buy food. From a general knowledge of social norms, he might also have inferred that, as a mother, she had a moral obligation to feed them. Thus, he might also have inferred that this is a situation in which two moral imperatives are brought into conflict. However, he seems to have shown no awareness of this at all, concluding simply that the woman should be arrested because "it is illegal not to go through the checkout line." Thus, his judgment might have stemmed from nothing more than a constrained and routinized concern with situational rule following.

Of course, more would need to be established to come to any firm conclusion in this case. For instance, if the conflict between these two moral imperatives were explicitly laid out for the young man, would he show more hesitancy in his judgment about what should be done? Still, the general lesson remains that just as knowing intent can matter for making appropriate first-order moral judgments (was the woman malicious or desperate?), so it matters for determining whether first-order moral judgments genuinely reflect any moral understanding on the part of the person making the judgments. The rule-following judgments and behavior of some individuals with autism may be too unreflective for that.

I have emphasized this aspect of autistic disorder—obsessive desire for sameness, hence a need for rules and routines—for two different reasons.

First, as I have already indicated, I do think some caution needs to be exercised in interpreting the "moral" behavior of individuals with autism. It may well be that their particular abnormalities give rise to a simulacrum of naive or innocent moral sensibility when in fact this sensibility can hardly be attributed at all. However, I don't wish to be entirely downbeat. It is certainly true, as Kennett points out, that the passion for rules and routines—hence for order in the world—already sets autistic individuals dramatically apart from psychopaths. They are prepared, as psychopaths are not, to discipline their own behavior and to judge others' behavior according to rules that they regard as universally binding. Moreover, it is difficult to say at what point having the mere simulacrum of moral sensibility shades into having a naive or innocent moral sensibility where this involves having some understanding of the point of the rules one follows (for instance, to avoid harming others). Equally, it is difficult to say at what point having a naive or innocent moral sensibility shades into the beginnings of a more sophisticated and genuinely autonomous moral sensibility where this would involve understanding such rules as answering to a principled moral sense of how things ought to be (for instance, as with the Amish teenagers (see note 5), that harming others is wrong in itself and hence to be avoided). Finally, since we know that autism is a spectrum disorder, with individuals varying widely in their abilities and disabilities, it should be no surprise to discover that individuals with autism vary significantly in the degree to which they are able to develop a genuine moral sensibility. So, despite the cautionary note I have sounded by emphasizing autistic rule-boundedness, I agree with Kennett that many high-functioning individuals do become autonomous moral agents; i.e., they become able and willing to govern their own behavior and to judge the behavior of others by reference to a deeper, more reflective consideration of the ends such behavior might be thought to serve.

Nevertheless, my second and larger point in emphasizing the predilection for rules and routines among individuals with autism is this: The need to impose order as a way of managing their environment predisposes high-functioning individuals with autism toward using their reason in a particular way. Specifically, it predisposes them toward discovering easy-to-follow principles behind whatever system of rules they find in place, even if those principles may be rather idiosyncratic from a nonautistic point of view. This point I think is nicely illustrated by Temple Grandin's reflections on her own rule-following behavior, which she sees as characteristically autistic:

For people with autism, rules are very important, because we concentrate intensely on how things are done. . . . Since I don't have any social intuition, I rely on pure logic, like an expert computer program, to guide my behavior. I categorize rules according to their logical importance. It is a complex algorithmic decision-making tree. There is a process of using my intellect and logical decision-making for every social decision. Emotion does not guide my decision; it is pure computing.

Learning a complex decision-making process is difficult. I had a strict moral upbringing, and I learned as a child that stealing, lying, and hurting other people were wrong. As I grew older I observed that it was all right to break certain rules but not others. I constructed a decision-making program for whether rules could be broken by classifying wrongdoing into three categories: "really bad," "sins of the system," and "illegal but not bad." Rules classified as really bad must never be broken. Stealing, destroying property, and injuring other people are in this category, and they were easy to understand. The "illegal but not bad" rules can often be broken with little consequence. Examples would be slight speeding on the freeway and illegal parking. The "sins of the system" category covers rules that have very stiff penalties for seemingly illogical reasons. Using my system has helped me negotiate every new situation I enter. (Grandin, 1995, pp. 103–104)

Grandin's "sins of the system" is a particularly interesting category, both for the kinds of norms or rules she classifies as such and for her attitude toward such rules. They tend to be rules having to do with social propriety or maintaining social order, and her attitude toward them is that these rules should be carefully observed even though she doesn't see the logic behind them. So why does she think they should be kept? The answer is probably mixed. At some level, she seems to manifest some concern for the social order as such, even though, as she often reports, it is also an order that she finds strangely alien. I discuss this possible motivation more in a later section. However, at another level, her motivation seems rather more expedient. For instance, she claims that in high school she regarded sex and smoking as the two greatest "sins of the system" and soon worked out "through careful observation and logic" that the teachers would give her considerable free rein as long as they were convinced that she would never engage in these prohibited activities (1995, pp. 102–103). In later life, her attitude toward sex and the social norms surrounding it has remained instrumentally cautious: "I still consider sex to be the biggest, most important 'sin of the system' . . . It has caused the downfall of many reputations and careers. . . . I've remained celibate because doing so helps me to avoid the many complicated social situations that are too difficult for me to handle" (p. 133). Thus, we see in much of Grandin's rule-following behavior a rationally driven response to her persisting need to

simplify, to order, to maintain clarity and control, even at the cost—if it is a cost—of avoiding what others consider to be morally loaded terrain (“carrying stiff penalties for seemingly illogical reasons”). Moreover, even in cases where she may evince a deeper understanding of why certain prohibitions exist, the need for clarity and control seems to play an important role in motivating her to toe a relatively “pure” moral line. For instance, with respect to lying she writes:

Autistic people tend to have difficulty lying because of the complex emotions involved in deception. I become extremely anxious when I have to tell a little white lie on the spur of the moment. To be able to tell the smallest fib, I have to rehearse it many times in my mind. I run video simulations of all the different things that the other person might ask. If the other person comes up with an unexpected question, I panic. Being deceptive while interacting with someone is extremely difficult unless I have fully rehearsed all the possible responses. Lying is very anxiety-provoking because it requires rapid interpretations of social cues to determine whether the other person is really being deceived. (Grandin, 1995, p. 135)

It may be premature to base any strong conclusions about the basic structure of autistic moral agency on these observations about their rule-following behavior and the motivation behind it. Still, they do suggest that we must be cautious in assuming too much commonality, or at least commonality at the wrong level, between individuals with autism and typically developed individuals. However, it still seems fair to ask with respect to autistic individuals if reverence for reason might not be the core moral motive in them?

For reasons I will come to in the next (and concluding) section, I hesitate to speak of a core moral motive, even in individuals with autism. Still, I do agree with Kennett that in comparison with the moral agency of typically developed individuals, autistic moral agency seems far less permeated by affect and more deeply governed by reason. As Grandin says repeatedly of herself, navigating in the social world is a “strictly logical process.” Now we might ask, why does reason speak with such force in autistic individuals? Why are they so prone to organize and judge their own and others’ behavior in terms of rules they are willing to treat as universally binding, even if they have no direct affective insight into the rationale for at least some of these rules? The answer I am suggesting is that individuals with autism have an unusual (arational) passion for order, and it is this passion for order that both motivates their rule-oriented behavior and encourages them to such virtuoso displays of reason in trying to enlarge their understanding of the kind of order that exists in the social world so that they might participate in it.⁸

If the passion for order is as dominant in individuals with autism as this proposal suggests, then we might expect to see it manifested in all sorts of ways over and above their rule-following behavior. Indeed, there is good evidence for this. For instance, there are many reports of unusual play behavior in very early childhood. Specifically, children with autism are noted for lining their toys up in rows, or treating all the objects they encounter in a particular way (e.g., trying to make them spin). As they grow older they often develop unusual interests that may be quite idiosyncratic but which nevertheless have a certain taxonomic or ordering quality in common. For instance, there are reports of individuals with autism becoming obsessed with timetables, bus routes, birth dates, door colors, and even types of vegetables. Happé gives the example of one young man who learned the name of every type of carrot, of which there are more than fifty, just to be able to name them; he had no other interest in carrots (Happé, 1994b, p. 37). Grandin herself manifests her own passion for order in a number of remarkable ways. For instance, when she was a high school student learning about entropy, she claims to have “hated the second law of thermodynamics because I believed that the universe *should* be orderly” (1995, p. 193). This led in turn to a “totally logical and scientific” belief in God as “an ordering force that was in everything” (p. 193). Moreover, she has made a hobby over many years of collecting “many articles about spontaneous order and pattern formation in nature” because “I want scientific proof that the universe is orderly” (p. 192). In sum, we see this passion for order manifested in many different ways among individuals with autism and at all cognitive levels. Furthermore, among those who are relatively high functioning, it should be no surprise to see a peculiar reverence for reason in them, since reason is a tool par excellence for discovering or imposing order in the world, especially in the social world, which we know they find especially challenging.

Let me now return to the questions I raised at the beginning of this section and draw some provisional conclusions. First, I do agree with Kennett that being a moral agent requires one to have certain agential capacities, in particular the capacity to control one’s impulses in the service of some larger ends (i.e., ends that trump one’s parochial and immediate interests), the capacity to find value in these larger ends, and the capacity to put one’s reason to work in the service of attaining such ends (minimally, through impulse control and more substantially through means-end reflection as well as reflection on the relative value of potentially competing ends). In my view, this makes reason, and the capacity to respond to the reasons one has, a prerequisite for moral agency. However, something

must explain what motivates one to respond to the reasons one has, and indeed to use reason to reflect on the relative values of one's ends. Here I think affect does play a critical role along the lines that Hume proposed: that is, certain ends become particularly salient for us because of our (possibly idiosyncratic) affective investment in them. Hence they become ends that compel our rational attention, giving reason a platform from which to speak.

My second conclusion, then, speaks against Kennett's Kantianism, as far as I understand it. Reverence for reason is not the core moral motive; indeed, I doubt it is much of a motive at all, at least on its own, as something pursued for its own sake. Rather, in my view, respect for reason (which in some individuals may rise to reverence) derives from the practical (so not necessarily consciously endorsed) recognition that it is one of the most useful tools we have for prioritizing and accomplishing whatever ends we find affectively salient or compelling, whether they be finding and imposing order in the world, promoting our own or others' interests, or whatever. Thus, affect must play a critical role in moral agency, i.e., affect that is something apart from mere reverence for reason.⁹ Moreover, *pace* Kennett, I take this rather Humean conclusion to be supported by a consideration of autistic moral agency. While autistic individuals may be lacking in the kind of empathetic attunement that provides the backbone of a typically developed moral agency, they do have a strong affective interest in living in the kind of world that is orderly, predictable, and indeed, respectful of individual space. They like clear boundaries and prefer social transactions that are aboveboard and explicit. Thus, it is no surprise that we find in them an interesting and substantial variety of (genuine) moral agency.

My third and final conclusion of this section is therefore a pluralist one: I say that it is thanks to the predominating affective concerns of autistic individuals that many of them are able to develop a genuine kind of moral agency. It is genuine so far as it goes beyond a mere predilection toward rule-following for rule-following's sake. However, it is a distinctive kind of moral agency as far as the affective profile that underpins autistic valuing of certain ends and that consequently ensures a respect for others and their modes of life is substantially different in many ways from the affective profile that characterizes a more typical form of moral agency. Here I agree with Kennett, this time against Hume, that sympathy is not the only possible source of moral concern. However, Kennett is wrong. I think to suggest that Humean sympathy must be replaced by some single fundamental source of moral concern that autistic individuals and normally

developed individuals share in common. Yes, they do share a certain agential capacity for responding to reason, a capacity that high-functioning individuals are particularly prone to cultivate. And Kennett may be quite right to suggest that psychopaths are seriously impaired in this regard. However, the agential capacity for responding to reason is rooted in the capacity for valuing certain ends, and valuing certain ends is fundamentally rooted in the depth and quality of one's affective life. Since the affective lives of autistic individuals are substantially different from the lives of normally developed individuals, we should expect to see differences in the sort of ends that are valued and in the priorities assigned to these ends.

This accounts, I think, for the difficulties sometimes evinced in giving a consistent assessment of autistic moral behavior relative to the norm. On the one hand, it is sometimes claimed that individuals with autism display a kind of moral purity or innocence in their interactions with others that approximates a sort of moral ideal; but, on the other hand, they can also display rigidity, insensitivity, and even callousness toward others that makes their behavior fall rather short of any ideal. In my view, this inconsistency is to be expected. The truth is that autistic moral behavior must always be a mixed bag relative to the norm since it is driven by affective concerns that are rather different from the norm. This is not to say there is no common ground in these different varieties of moral agency, as I will next try to show. Rather, it is to say that a family resemblance in surface behavior need not imply the existence of identical or even substantially similar cognitive and/or affective profiles. We have already learned this lesson in the social cognitive domain as a consequence of trying to explain how high-functioning individuals with autism are sometimes able to reason about others' mental states despite impairments in their so-called theory of mind abilities. Perhaps it is time to learn this same lesson and explore its implications in the domain of moral psychology.

Varieties of Moral Agency: Speculative Reflections

I began this chapter by asking what makes us moral creatures. How do we come to make moral judgments and act from specifically moral motivations? What are our moral concerns? As the foregoing discussion ought to make plain, I think some general answers can be given to these questions, answers that acknowledge the importance of reason in our moral lives but which nevertheless give special attention to the central role of affect. Specifically, I claim in a broadly Humean way that we human beings are moral beings—and indeed the kind of moral beings we are—because of our

affective natures. Our moral intuitions are generally grounded in a range of emotions that are part of the way we experience the world. Moreover, we develop our capacity for heeding the dictates of reason just because the ends for which our reason speaks are affectively charged; hence they become the ends we are able to value over our immediate or parochial interests. I will not say more in defense of this general position, for now my interest is in exploring, in a purely speculative way, the proposal I have made with respect to autistic moral agency: that the range of emotions in which their moral intuitions are grounded departs substantially from what might be considered the typical or normal range, producing in them a genuinely distinctive variety of moral agency. As I see it, this proposal faces a conceptual challenge that I would like to address briefly in this concluding section: namely, in what sense could the affective profile of autistic individuals be that different from that of typically developing individuals while still supporting what is recognizably a genuine, albeit distinctive, variety of moral agency? In other words, how can there be enough commonalities among the differences, and enough differences among the commonalities, to give rise to this possibility?

To sketch an answer to this question, I begin by noting that a preoccupation with empathy in the domain of moral psychology can narrow our focus unduly when it comes to identifying the range of affective states that normally underlie our moral lives. Even when we understand empathy to be not an emotion in itself, but rather a disposition to be affectively attuned to, and even appropriately responsive to, another's affective states, it often carries a connotation of being compassionate, caring, or concerned for the well-being of another. However, as Jonathan Haidt has argued, not all morally relevant emotions can be understood in these terms. As he says, "there is more to morality than altruism and niceness. Emotions that motivate helping behavior are easy to label as moral emotions, but emotions that lead to ostracism, shaming and murderous vengeance are no less a part of our moral nature" (Haidt 2003b, p. 855). To arrive at this more inclusive understanding of the moral emotions, Haidt recommends that we take a functional approach to their identification. That is, even though we normally classify emotions such as anger, fear, sadness, joy, and so on according to their distinctive facial expressions, physiological changes, and phenomenological tone, we count certain manifestations of these as moral just in case they have particular kinds of eliciting conditions and give rise to particular kinds of action tendencies.

The rationale for this approach can be made evident once we follow Haidt in regarding all emotions as action-priming "responses to perceived

changes, threats or opportunities in the world" (Haidt 2003b, p. 853). Of course, as he points out, many of these responses are concerned with our own self-interests narrowly conceived. However, we are perhaps unique among other species in expending a remarkable portion of our emotional energy reacting to events that have no direct impact on such interests, but are rather seen to affect the overall shape and structure of our (social) world and the other creatures (especially other people) in it. The philosopher P. F. Strawson is well known for making a similar point, observing that many of our "reactive attitudes"—a subset of emotions provoked by the activities of responsible agents—are felt precisely on behalf of others; i.e., they are felt in consequence of our perceiving someone else to be harmed or benefited by another person, even if we are not directly harmed or benefited ourselves. Strawson calls such reactive attitudes "impersonal," "generalized," or "vicarious," noting that there are particular emotions, such as indignation, that are particularly apt for being provoked in this way (Strawson, 1974). More interestingly still, we have many "self-reactive" attitudes by which we approve or disapprove of our own behavior as far as we regard that behavior as producing benefits and harms to others. Here pride, shame, and guilt are prime examples. In line with these observations, Haidt proposes a general scheme in which we classify our emotional reactions as moral to the degree that (1) they have "disinterested elicitors"; i.e., they are provoked by events touching concerns that reach beyond our narrow self-interest and (2) they have disinterested "action tendencies" (Haidt calls these "prosocial"); i.e., they prime us (motivationally and cognitively) to act in ways that benefit others or that uphold or benefit structures that we value, such as the "social order."

There is much to be said in defense of Haidt's functional characterization of the moral emotions. It is intuitively plausible, theoretically well motivated, and conceptually attractive in its simplicity and generativity. My only complaint, if it is a complaint, concerns Haidt's (understandable) tendency to focus exclusively on our more socially oriented interests. Thus, by way of a preliminary definition, he suggests that moral emotions are "those emotions that are linked to the interests or welfare of society as a whole or at least of persons other than the judge or agent" (Haidt, 2003b, p. 853). This is fine as far it goes. I agree with Haidt that in general we find in human beings two primary—and I would say distinct—spheres of disinterested concern.¹⁰ These are a concern with others' well-being and a concern with the structure and maintenance of the social order, giving particular attention to how individuals find and occupy appropriate social roles. I call these distinct spheres of disinterested concern because I would

argue they are rooted in quite distinct affective-cognitive systems, the first being the attachment system and the second being a system devoted to the production and distribution of social goods.

This first system has been much discussed in the developmental literature, and it makes sense from an evolutionary point of view if only because *human neonates need to be strongly attached to particular significant others to ensure their own survival and development.* As for the second system, it too makes sense from an evolutionary perspective since, like other social mammals, we are highly dependent on structured cooperative relations for acquiring and distributing resources. Thus, we are programmed to care about how those relations are maintained; we are programmed to care about how we and others fulfill our social roles, as well as what our particular social roles with their own rights and responsibilities should be. Still, important and predominating as these two spheres of concern are, I think there is yet a third sphere of disinterested concern that is also most likely rooted in a distinct cognitive-affective system. I tentatively label it (3) a concern with "cosmic" structure and position.

Why think there is such a distinct sphere of disinterested concern? To begin at the purely behavioral level, it seems clear that human beings are uniquely preoccupied with questions about the meaning of life, about the origin and fate of the universe, about our place in the great scheme of things, and about whether or not there is any great scheme of things at all. Moreover, these are not simply intellectual preoccupations. We care passionately about there being order in the universe, about there being some entity or entities—for instance, the Judeo-Christian God—that gives meaning and shape to it all, indeed, finding in such passions remarkable reservoirs of faith for systems of belief that otherwise have very little evidential support. Even among that small minority of individuals whose intellectual predilections and/or training prohibit any comfortable acceptance of belief on faith, many profess feelings of awe or wonder at the beauty science reveals in the ordering laws and patterns at all levels of nature, and even of deep contentment in the recognition that we human beings have our own place in all of that.

Why should the existence of such cosmic order matter to us so deeply? Why do we find it cognitively and aesthetically so appealing? Why does it inspire such reverence? And why do we feel a deep-seated need to secure our own place in it? My answer to these questions must be incomplete and provisional. However, I suggest that these affectively laden concerns are at least partially rooted in pattern-seeking cognitive machinery that is uniquely well developed in *Homo sapiens* and which is dedicated to impos-

ing order and meaning on our interactions with the physical world across time, making it seem a more stable place to us and locating for us a stable place within it. Once these points of reference are in place, we are motivationally primed to engage in long-term planning that leads to better success in navigating our environment. If this extremely sketchy account is on the right track, then at least there is an obvious evolutionary explanation for why a disinterested and deeply felt concern for cosmic structure and position is present in our species. Even if it isn't, the fact that we have such cosmically oriented affective concerns cannot be denied; and it becomes an interesting open question as to why we should have them, given that they seem unrelated to either our concern for the well-being of others or our concern for the social order.

Here, then, in a nutshell is my speculative proposal about the different varieties of (human) moral agency. I begin by summarizing what I think they share in common. To wit: all forms of human moral agency are rooted in affect. We are the kind of moral beings we are because we have powerful emotional reactions to certain kinds of events or situations; namely, events or situations that touch upon various disinterested concerns. Furthermore, in all human beings there are three distinct varieties of disinterested concern, rooted, I suggest, in distinct cognitive-affective systems: (1) a concern for the well-being of others, (2) a concern with social structure and social position, and (3) a concern with cosmic structure and cosmic position. Given these concerns, various events or situations will provoke different kinds of emotional responses, priming us to take different kinds of action. Sometimes, of course, our concerns will lead to emotional responses that are mutually reinforcing. For instance, I may feel indignant about someone causing another person harm both because it is socially disruptive and because it compromises the other's well-being. However, this won't always be the case. For instance, given my concern for the well-being of others, I may be inhibited in causing someone else distress. Yet, given my concern for social structure and position, I may feel angry with that person for offending against a social norm, provoking a desire to punish them and so cause them distress. In other words, these different spheres of concern can lead to emotional responses that pull in different, sometimes even conflicting directions. How we resolve such conflicts may well depend on which kind of concern is most dominant in us.

So here is my first suggestion: Moderately different varieties of moral agency can emerge as a consequence of how these three spheres of disinterested concern develop and interact in a given person, varying according to individual differences as well as under the sway of different cultural

influences. Still, in typically developing individuals, we can expect to see a close family resemblance among these varieties of moral agency, for it seems to be a near-universal feature of the human affective profile that we are very much dominated by our concern for maintaining social order, hence for policing the ways individuals succeed or fail in playing their appropriate social roles. I say this because we have developed specialized skills for operating in the social world—our much-vaunted mindreading abilities—and we have a well-developed range of emotional responses that are very much adapted to the intricate patterns of our social interactions. Thus, we have other-condemning emotions such as anger, contempt, and disgust; self-condemning emotions such as shame, embarrassment, and guilt; other-praising emotions such as admiration, humility, and respect; and finally self-praising emotions such as pride and self-respect. Of course these emotional responses may be moderated in various ways by our compassion or sympathy for others, since we must, by many accounts, have rather well-developed empathetic capacities if we are to develop the advanced social-cognitive skills that support our intricate social interactions. Thus, we should expect to find in typically developing individuals a fairly well-entrenched concern for others' well-being, particularly for those with whom they are personally connected. Of course, the concern for cosmic order may be rather well developed too, although this seems to be the sphere of concern most deeply affected in typically developing individuals by cultural (including educational) influences. In any case, I think it is the dominance of our concern for social place and the extensive range of emotional responses we have developed as a consequence that explain the familiar paradox with which I began this chapter: that our capacity for cruelty as much as our capacity for kindness is rooted in our moral being.

I turn now to the question of autistic moral agency. How can it be such a distinctive variety of moral agency while still being distinctively human? My proposal is that what makes autistic moral agency distinctively human is that, just as with typically developing individuals, these three spheres of disinterested concern are operative in individuals with autism: concern with others' well-being, concern with social order, and concern with cosmic order.

This claim may seem surprising. After all, with regard to the first concern, a marked lack of empathy has traditionally been cited as a diagnostic feature in autistic spectrum disorder. Children with autism do not seem to tune into other people at all, seemingly even from the earliest stages of postnatal development. And yet, despite this fact, various studies show

that a significant portion of children with autism do manifest some (maybe unusual) form of attachment behavior (Capps, Sigman, & Mundy, 1994; Dissanayake & Crossley, 1996, 1997; Rogers, Ozonoff, & Maslin-Cole, 1991; Shapiro, Sherman, Calamari, & Koch, 1987; Sigman & Mundy, 1989; Sigman & Ungeler, 1984). Furthermore, although their reactions are typically muted compared with those of normally developing children, there are children with autism who are able to recognize that others are in distress, and some even offer gestures of comfort (Bacon, Fein, Morris, Waterhouse, & Allen, 1998; Dissanayake, Sigman, & Kasari, 1996; Sigman, Kasari, Kwon, & Yirmiya, 1992; Haviland, Walker-Andrews, Huffman, Toci, & Alton, 1996). Based on his data, Blair has argued that children with autism do make the moral-conventional distinction, and this is because they are sensitive—unlike psychopaths—to the distress of others (R.J.R. Blair, 1996). For instance, they appear to show heightened autonomic response to pictures of distressed faces compared with pictures of neutral faces (R.J.R. Blair, 1999). Thus, some basic concern for the well-being of others is independent, Blair suggests, of advanced mindreading skills (R.J.R. Blair, 1996). However, given their very deep impairments in tuning into others and so developing advanced mindreading skills, it is no surprise that this basic concern with the well-being of others should remain fairly basic.

It is even less surprising that autistic individuals have a very unelaborated emotional repertoire relating to the most dominant of our concerns; that is, a concern for the social order and one's place within it. Nevertheless, I suggest that this concern is operative, at least at a basic level, in individuals with autism. Thus, we see many high-functioning individuals express a desire to fit in despite their rather heartbreaking awareness of their own inability to do so. "Passing for normal" and so observing the forms of social life is something with which they are greatly preoccupied.

Finally, we come to the concern for cosmic order. Here I think we find a sphere of concern that is underpinned by a relatively intact cognitive-affective system. Thus, it is this sphere of concern that dominates in autistic moral agency, and dramatically so. In consequence of this, we see the emergence of an entirely distinctive style of human moral agency, where the usual order of dominance among spheres of disinterested concern is completely inverted. Indeed, it is more than inverted; concern for social place and, to a much lesser extent, concern for the well-being of others have only the crudest of roles to play in shaping the emotional responses of autistic individuals to the kinds of situations that elicit such responses.

Because these are speculative proposals, I will not elaborate in any more detail. However, I want to close with a word about psychopathy. As mentioned in note 3, *DSM-III* replaced the term "psychopathy" with "antisocial personality disorder," a term that is still used in *DSM-IV* and *DSM-IV-TR* to refer to this disorder. Researchers and clinicians have questioned the validity of the diagnostic criteria associated with this change of label, claiming that too much emphasis has been placed on behavioral traits over far more indicative personality traits (R.J.R. Blair, Blair, Mitchell, & Reschardt, 2005; Hare, 1996b). In addition to this complaint, I find the new label somewhat ironic. On the one hand, I agree with Kennett that individuals with psychopathy are most likely seriously impaired in their capacity to use reason at all. However, in my view, this is because they have limited capacities for making any affective investment in ends that transcend their immediate and parochial interests. In other words, the three spheres of disinterested concern that are normally operative in human beings do not seem to be operative in them, owing to an overall flattening in the affective tone of their cognitive operations.

Now for a wild speculation: If concern for social place is strongly dominant in us and is supported by specialized skills in mindreading, then one might expect to see some faint semblance of at least this concern operating in individuals with psychopathy, especially given the fact that they seem to be relatively good at mindreading. Indeed, I think this may be the case. Individuals with psychopathy do show some concern with the social world, at least as far as that extends to getting the better of others whom they imagine to be trying to get the better of them. It is of course a seriously distorted concern with social position and social order, not tempered in the least by other sorts of disinterested concerns (for others' well-being, for cosmic structure and position) or even by a well-elaborated social emotional repertoire. Nevertheless, there is something characteristically human—indeed, something not unrelated to normal moral agency—that explains the psychopath's quest for dominance in the social world. Perhaps this also explains why they often do a much better job than autistic individuals at passing for normal (Babiak, 1995; Hare, 1996b).

Notes

1. In Kohlberg's own words: "We are claiming . . . that the moral force in personality is cognitive. Affective forces are involved in moral decisions, but affect is neither moral nor immoral. When the affective arousal is channelled into moral directions, it is moral; when it is not so channelled, it is not. The moral channeling mechanisms themselves are cognitive" (Kohlberg, 1971, pp. 230–231).

2. A word of caution: The term "empathy" is used in a wide variety of ways in the philosophical and psychological literature, so its meaning cannot be assumed. Often it is used to identify a particular other-regarding or other-directed emotion. Candidates are (1) concern for another's well-being, although I prefer the term "compassion" for this, and (2) distress at another's distress, more clearly designated by the acronym DAAD. However, the term has been better used, I think, not for a single emotion, but rather for an other-regarding disposition toward feeling or being affectively moved by the emotional state of another, whether that state be distress, joy, anger, or whatever (Eisenberg, 1991). This other-regarding disposition may be closer to what Hume meant by "sympathy," although certain passages suggest he might have thought sympathy in the sense of a particular emotion—i.e., compassion or concern—is necessary for having the dispositional capacity to respond in an affectively appropriate way to the feelings of another: "No quality of human nature is more remarkable, both in itself and in its consequences, than that propensity we have to sympathise with others, and to receive by communication their inclinations and sentiments, however different from and even contrary to our own" (Hume, 1740/1978, p. 743).

3. Psychopaths are now subsumed under what many researchers consider to be a more general (and less valid) diagnostic category: "antisocial personality disorder" or ASPD (see, for instance, *DSM-IV-TR*, 2000, p. 702). In the words of Robert Hare, "most psychopaths (with the exception of those who somehow manage to plow their way through life without coming into formal or prolonged contact with the criminal justice system) meet the criteria for ASPD, but most individuals with ASPD are not psychopaths" (Hare, 1996b). ASPD is diagnosed primarily on the basis of behavioral criteria, e.g. criminal behavior, so it is no surprise that ASPD is common in criminal populations. Psychopathy, on the other hand, has been more narrowly defined in terms of measurable interpersonal and affective characteristics, i.e., sustained personality traits, that will often produce criminal behavior. Such characteristics include "egocentricity, deceit, shallow affect, manipulativeness, selfishness, and lack of empathy, guilt or remorse" (Hare, 1996b). While it is no surprise to see such traits in the criminal population, Hare cautions that not all psychopaths will engage in obviously criminal behavior: "[p]sychopaths have little difficulty infiltrating the domains of business, politics, law enforcement, government, academia and other social structures" (Hare, 1996b, p. 40). Since my aim is to focus on this population, I will continue to use the term "psychopath" in this chapter, rather than "individuals with ASPD."

4. See the chapters by Kiehl and Kennett and Fine in this volume.

5. A particularly nice example of the authority dependence versus authority independence of different sorts of transgressions can be found in Nucci's study of Amish teenagers: 100 percent of those tested claimed that if God made no rule against working on Sunday, it would not be wrong to work on Sunday. By contrast, more than 80 percent claimed that if God had made no rule against hitting someone, it

would still be wrong to hit (Nucci, 1986). This brings to mind Socrates' Euthyphro question: Is an act just because the gods love it, or do the gods love it because it is just? Apparently for Amish teenagers, the answer depends on the nature of the act.

6. It is interesting that Blair's work on psychopaths shows that while they fail to make a significant moral-conventional distinction, they tend to process all transgressions as moral, at least as far as the criterion of authority jurisdiction is used. That is, the psychopaths tested tended to claim that both moral and conventional transgressions would be wrong independently of whether or not the acts in question were prohibited by someone in authority. However, when they were asked to explain why the acts would be wrong, the test subjects made significantly less reference to others' welfare or the existence of potential victims even in cases (typically identified as moral transgressions) where harm to the victim was clear. Hence Blair concludes that the assimilation of conventional to moral transgressions on the criterion of authority jurisdiction might well be an artifact of the population tested. That is, all the test subjects were "incarcerated and presumably motivated to be released. All wished to demonstrate that the treatments that they were receiving were effective. They would therefore be motivated to show that they had learned the rules of society" (R.J.R. Blair, 1995, p. 23) and presumably they would give assurances that they were prepared to abide by the rules, come what may. In any case, as Blair rightly points out, these test subjects were not able to make the distinction normally, in marked contrast to the control group (nonpsychopathic fellow inmates).

7. Actually, there is evidence that some autistic children do experience "empathy" in the less complicated sense of feeling distress at another's distress. I return to this issue later.

8. Kennett also takes note of the passion for order found in many autistic individuals (Kennett 2002, pp. 350–351), but seems to tie this to their reverence for reason. However, even if one places reverence for reason on the appetitive side of human psychology (see note 9), it seems to me that the passion for order is something quite distinct. Rationality may involve a preoccupation with certain kinds of order (e.g., consistency in one's beliefs), but there may be many kinds of order (e.g., lining up ducks in a row or ensuring that one takes a walk at precisely the same time every day) that serve no rational purpose at all. There is ample empirical evidence that autistic individuals care about order in this larger extrarational sense.

9. One complicating feature in my disagreement with Kennett is this: Contrary to the standard, and perhaps caricatured, contrast between Kant and Hume, Kennett maintains that moral feeling plays an important role in Kant's account of moral psychology. To wit, that we would not be moral creatures were we not affectively moved to respond to the dictates of reason. Hence, it seems that, according to Kennett's Kant, what motivates us to respond to the reasons there are is the sui

generis desire to think and act rationally. If this is an accurate representation of Kennett's (and/or Kant's) view of how human beings are psychologically structured, I am happy to be somewhat concessive. Indeed, Kennett cites some interesting empirical evidence on cognitive dissonance to support the claim that we are naturally and normally endowed to like acting in accord with reason and to dislike acting against it (2002, p. 354). (Rationalization, as Kennett points out, is a handy way to overcome our affective distaste for contravening the dictates of reason.) Still, my claim is that while the desire to think and act rationally may be deeply rooted in our (normal) human nature, I do not think it is sufficient on its own to account for moral thought and action, even in individuals with autism.

This is not to say I am taking a stand on the conceptual issue of whether it is possible for there to be creatures endowed simply with the desire to think and act rationally that are *ipso facto* moral creatures. Philosophically, I think this is a difficult thesis to defend, but my claim here is more modest. If you like, it is to deny a certain sort of existence proof: that autistic individuals exemplify such a type. It may be that autistic individuals do evince a stronger desire to think and act rationally than is normally found among typically developing human beings; there is some evidence of this. It may be that in autistic individuals the desire to be rational is experienced less as a means to other goals and more as an end in itself than in typically developing individuals; I think there is less evidence for this. Still, the question remains: What kind of affectively loaded interests do autistic individuals have—apart from being rational—that play a significant, indeed critical, role in the development of their moral sensibility? I have identified their passion for order as one such interest, but in fact I think this is only part of the story. I say more about the affectively laden interests of autistic individuals in the concluding section.

10. I follow Haidt and indeed many others in the literature in using the term "disinterested concern" to talk about interests that are not narrowly concerned with the self, either in terms of their focus or in terms of their playing into a calculus of costs and benefits accruing exclusively to oneself. Here the term is not meant to imply what it often does, namely, an emotionally neutral preoccupation, perhaps supported by reason alone.

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 apology 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 Malbom), 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 Malbom 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 *Ashta*, 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 Malbon 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. Maiborn, 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, Maiborn 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 Maiborn 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 Maiborn'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 sophistication 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 Fritth 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 Fritth, 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. Malbom 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.