Conversation, Responsibility, and Autism Spectrum Disorder[1]

Section 1: Introduction

In his recent book, *Conversation and Responsibility*, Michael McKenna presents a compelling new theory of moral responsibility which is modeled on an analogy with a linguistic conversation. Picking up on Gary Watson's suggestion that our practices of holding others responsible serve a fundamentally expressive function,[2] McKenna constructs a theory in which this communicative nature of responsible agency and its attendant practices are brought to the fore. While impressive in both its scope and nuance, this conversational theory is open to an important objection from the moral psychology and neuroscience literature. Though the theory may successfully characterize a great many moral agents, it cannot, I argue, account for at least one group of agents, namely, those with autism spectrum disorder (ASD)[3]. High-functioning autistic individuals, as well as those with Asperger Syndrome (AS),[4] represent a class of individuals which serves as a genuine counterexample to the conversational theory of responsibility, and examining the

[2] Watson, "Responsibility and the Limits of Evil"
[3] There is an ongoing debate regarding the use of the phrase "individuals with ASD," as opposed to "autistic individuals," to which philosophers ought to be sensitive. Some find the former phrase objectionable because it seems to imply that autism is a type of disease or disorder that ought to be cured or eradicated (this view is often associated with the Neurodiversity movement). Others find the latter phrase objectionable because it seems to suggest that one's identity consists in being autistic rather than in other features of the person him or herself. In this paper, I will use the phrase "autistic individuals" as this seems to capture the more prevalent preference of those in the autistic community. However, I by no means intend for this phrase to imply any sort of judgment regarding the value of autistic persons or of what it is like to be autistic more generally.
[4] For the sake of convenience, I will often refer to autistic individuals, *simpliciter*. However this should be understood to refer to high-functioning individuals (more on this distinction below) and to include individuals with AS as well.

distinctive moral psychology of such individuals can help to show precisely where the theory goes wrong.

In what follows, I will attempt to spell out the nature of the problem that autistic individuals pose for McKenna's theory. In section 2, I will offer an analysis of McKenna's view. In section 3, I will present a short prelude to the challenge that I will ultimately pose for his theory. In section 4, I will attempt to offer a partial profile of autistic moral psychology in an effort to highlight the morally salient features of that psychology which may be relevant to the conversational theory of responsibility. In section 5, I will argue that these moral psychological features which characterize many autistic individuals present a considerable problem for the conversational theory because they suggest that there are individuals in the world who meet all of the conditions for responsible agency that the theory lays out but who are nevertheless not responsible agents. Attending to the moral psychology of such individuals will, I think, help shed light on an important gap in the conversational theory.

Section 2: Understanding the Conversational Theory of Responsibility

There are several important features of the conversational theory that McKenna ably defends. Understanding these features and how they fit together will be the focus of this section. A complete analysis of the theory would require far more attention than I can give it here, and so, my goal in this section will simply be to give the reader a general sense of the theory where a detailed treatment is not necessary.

As I see it, McKenna's view consists of three theses. The first is a metaphysical claim about the relationship between being responsible and holding responsible. The second is a Strawsonian quality of will thesis, and the third is the claim that we should understand

responsible agency as being modeled on an analogy with a linguistic conversation. I will discuss each of these in turn.

2.1: The Modest Metaphysical Thesis

McKenna's first substantive claim is one about the relationship between being responsible and holding responsible, and in defending it, he takes up the debate between those who, on the one hand, maintain that holding others morally responsible is metaphysically more basic than being responsible – the normative interpretation – and those who, on the other hand, hold that there are facts about moral responsibility which are metaphysically prior to facts about holding responsible – the extreme metaphysical interpretation. Those who take the normative interpretation claim that our practices of holding others responsible take explanatory priority over the facts that obtain with respect to any given agent and are metaphysically more basic in grounding morally responsible agency. So, those who take this broadly Strawsonian[5] view of responsibility claim that the facts about the appropriateness of holding responsible and the norms that govern the practices of doing so serve to fix the relevant facts about being morally responsible. Thus, the norms involved in holding others responsible – or the conditions that make it appropriate to hold others responsible – are metaphysically more basic than any independent facts about responsible agents.

In contrast with this, those who endorse what McKenna refers to as the "extreme metaphysical interpretation" take the converse of the above approach. The extreme metaphysical interpretation says that there is a set of facts about responsible agency, and these facts are what determine our practices of holding one another responsible. Or, to put

---

[5] See Strawson, "Freedom and Resentment."

it another way, our practices of holding others responsible simply track the facts of responsible agency. On this view, then, being responsible is wholly independent of the norms which govern our practices of holding others responsible. In other words, being responsible is metaphysically more basic than both the practices involving holding responsible and the norms which govern such practices. As one proponent of this view, Michael Zimmerman, writes, "Someone is blameworthy … if it is correct, or true to the facts, to judge that there is a 'debit' in his 'ledger' (etc.). It is important to note that, in the context of inward moral praise or blame, worthiness of such praise or blame is a strictly nonmoral type of worthiness; it is a matter of the truth or accuracy of judgments."[6]

Rather than taking sides in this debate, McKenna attempts to argue for a compromise position which he calls the "modest metaphysical interpretation." He claims that, "Holding responsible should in the first place answer to the facts about what it is to be responsible. Yet … I contend … that being morally responsible presupposes considerations employed from the standpoint of holding responsible."[7] In other words, McKenna sees the facts about an agent's being responsible and the practices of holding her responsible as interdependent. As a result, he bases his theory around the notion that being responsible and holding responsible are fundamentally integrated with one another. Or, as McKenna puts it, "the metaphysical thesis proposed is that there is an irreducible relation of interdependence between being and (pertinent norms pertaining to) holding responsible."[8] So, according to the modest metaphysical interpretation neither being responsible nor holding responsible is metaphysically prior to the other. Instead, McKenna proposes that

---

[6] Zimmerman, *An Essay on Moral Responsibility,* pg. 38. As quoted in McKenna, *Conversation and Responsibility*, pg. 44
[7] McKenna *Conversation and Responsibility,* pg. 50
[8] Ibid. pg. 54.

we understand being responsible by reference to the quality of an agent's will and that we then understand the nature of quality of will in virtue of the capacities that an agent might have relative to a kind of moral responsibility conversation that is analogous to a successful linguistic conversation between competent speakers.

2.2: Quality of Will

In order to better understand McKenna's modest metaphysical interpretation, it will be important to understand the role that quality of will plays in his work. He proposes introducing a quality of will thesis which states, "Being morally responsible for an action is to be settled in terms of the moral quality of the will with which an agent acts."[9] However, given his modest metaphysical interpretation, this is to be partially understood with respect to the practices of holding responsible. Thus, he writes, "the moral quality of an agent's will, as well as a person's standing as a morally responsible agent, is dependent upon that agent's appreciation of the expectations of due regard for others as revealed in the practices of those holding responsible."[10] Given this, McKenna attempts to reorient the classical Strawsonian position around this quality of will thesis so that being responsible is dependent upon acting with a certain quality of will and being morally blameworthy is dependent upon how that will is revealed in an agent's actions. He then gives an account of the reactive moral emotions that is sensitive to this reorientation. On his view, the morally reactive emotions are responses to the quality of an agent's will when she acts, and this means that they are fundamentally cognitive insofar as they are directed at a propositional object such as the belief that one has acted with an objectionable quality of will. Moreover, McKenna argues that the reactive attitudes are best understood in terms of their overt

---

[9] Ibid. pg. 58.
[10] Ibid.

manifestation rather than any private experience of them. This means that in order for an individual to understand the reactive emotion that he or she experiences, he or she must understand what sorts of practices would be fitting manifestations of the experienced emotion.

So, on McKenna's view, what is of primary importance for determinations of responsible agency is a fact about the agent in question, namely the quality of the agent's will. Because of this, reactive attitudes are responses to this quality of will which are to be understood in cognitive terms and by reference to their public manifestations. It follows from this, McKenna argues, that agents who are exempted from being objects of the reactive emotions are those who cannot understand the practices of holding responsible when these practices are directed toward the agent in response to her being perceived as evincing an objectionable quality of will. Moreover, if the agent cannot understand these practices, then she cannot engage in them either, and this notion leads naturally to a discussion of the second portion of McKenna's modest metaphysical interpretation to which I will now turn.

2.3: The Moral Responsibility Exchange

On McKenna's view, the ability to hold others responsible is necessary in order for an agent to be responsible. In justifying this thesis, McKenna appeals to Paul Russell's notion of "moral sense." For Russell, the condition of moral sense is just the ability to "feel and understand moral sentiments or reactive attitudes,"[11] and, he argues, it is a necessary condition for morally responsible agency. "Agents who lack moral sense," Russell writes, "are missing something that is vitally important and that 'normal' agents do and must

---

[11] Russell, "Condition of Moral Sense," pg. 293.

possess. More specifically, when an agent lacks moral sense, her *sensitivity* to moral considerations is diminished and her motivation to be guided by these considerations is impoverished and limited."[12] Taking up this line of thought, McKenna claims that agents who are unable to hold other agents responsible by feeling and understanding the moral emotions lose access to a specific type of reason. One example of the reasons he has in mind are second-personal reasons like those proposed by Stephen Darwall. Of these, Darwall writes, "A second-personal reason is one whose validity depends on presupposed authority and accountability relations between persons and, therefore, on the possibility of the reason's being addressed person-to-person."[13] In further discussing the second-personal nature of responsibility, Darwall says, "I claim that reactive attitudes are always implicitly second-personal and that they therefore invariably carry presuppositions of second-personal address about the competence and authority of the individuals who are their targets, as well as about those who have them."[14] So, by making use of Russell's notion of moral sense, and Darwall's treatment of the second-personal features of responsible agency, McKenna argues that being responsible is dependent on an agent's ability to hold others responsible because without this ability an agent would lack the capacity to recognize and respond to second personal reasons as these reasons require an ability to hold others to second-personal demands. To this end, McKenna writes,

> If Darwall is correct about the very existence, nature, and pervasiveness of these reasons, and to my mind he is, we have a straightforward source of support for my claim about moral responsibility's dependence on the nature of holding morally responsible. The practices by which others hold one morally responsible are themselves expressions of demands that as a competent agent one must be able to grasp and treat as reasons that apply to one. In the absence of this ability, a person

---

[12] Ibid, pg. 294.
[13] Darwall, *The Second Person Standpoint*, pg. 8.
[14] Ibid. pg. 67.

would be unable to recognize and respond to a vast array of reasons presented to morally responsible agents.[15]

Given all of the above, McKenna proposes a theory of moral responsibility which gives full treatment to the communicative nature of our moral responsibility practices. In doing so, he builds a theory which presents responsibility as modeled on an analogy with a linguistic conversation saying,

> By acting as she does, the morally responsible agent *opens up the possibility of a conversation* about the moral value of her action, and most notably, what it reveals about the quality of her will. This initial contribution provides a basis for members of a moral community responding to her by holding her morally responsible, thereby engaging in a dialogue with her. Given the unfolding conversation, it is now the agent's place – her conversational role – to extend the conversation by offering some account of her conduct either by appeal to some excusing or justifying consideration or instead by way of an acknowledgment of a wrong done, perhaps an apology offered.[16]

So, there are three stages to the conversation, or, as he calls it, the "moral responsibility exchange." The first is the stage of moral contribution in which the responsible agent makes an opening salvo, followed by the stage of moral address in which her peers hold her responsible, and, finally, the stage of moral account in which the responsible agent gives some account of her actions.

In order to add legitimacy to the conversational model, McKenna proposes a further feature of the analogy. One obvious and important aspect of a linguistic conversation is that the speaker who makes an utterance typically intends for that utterance to be understood by her audience, that is, she intends her utterance to have *meaning*. In the same way, McKenna claims, "a morally responsible agent acts with the knowledge that her conduct is always a *potential* object of interpretation by members of the (or a) moral community. She

---

[15] McKenna, *Conversation and Responsibility,* pg. 84.
[16] Ibid. pg. 88-89.

therefore is able to understand her own actions as having meaning."[17] And that meaning, on this view, is found in what is revealed about the agent's quality of will through her actions. In summary, McKenna offers the following description of the theory he is proposing:

> [W]hen a morally responsible agent acts, she understands the standpoint of holding morally responsible, and so understands that others in a moral community are liable to take her actions as indicative of the quality of her will. In acting, she must be able to appreciate how her conduct could be interpreted as reflective of the quality of will with which she does act. And she must be able to adjust her behavior in ways sensitive to these considerations – or so I have argued. This is irrespective of whether as a matter of fact she allows those considerations to have any bearing upon her intentions or whether these matters even occur to her when she does act. *A morally responsible agent must be able to act as if her actions were morally significant bearers of meaning among communicating agents. And because of this, her actions thereby take on meaning merely by virtue of her being such an agent.*[18]

In short, McKenna's view is characterized by three basic theses. It is one in which being responsible and holding responsible are held to be interdependent rather than either having metaphysical priority. It is oriented around a quality of will thesis so that any judgment of responsibility must be made in reference to the quality of an agent's will. Most importantly it highlights the expressive nature of responsibility and its attendant practices by assigning a great deal of importance to the ability of an agent to take part in the moral responsibility exchange. Having understood this, we are now in a position to understand the challenge that autism spectrum disorder and its distinctive moral psychology poses for such a theory.

Section 3: The 'Functional' Moral Agent

Before discussing some important features of the moral psychology of ASD, I would like to pause to offer a prelude to the problem that ASD ultimately poses for the

---

[17] Ibid. pg. 94.
[18] Ibid. pg. 99 (emphasis added).

conversational theory. Given that an agent's status as morally responsible, on McKenna's view, is determined by the agent's ability to engage in the robust sort of moral responsibility exchange that he describes, it will be a problem for the theory if there are agents who can converse in the relevant ways but who are nevertheless not responsible. I'll refer to such agents as 'functional' agents.

To better grasp what a functional agent might look like, consider first the linguistic analog of such an agent:

> **The Functional Speaker:** Suppose Sandra is a speaker in a foreign land who has been forced to acquire the language of the indigenous speakers without the help of any formal training. She observes the people around her and is able to gain a functional fluency in the native language. She understands and is able to make use of a variety of meaningful statements. She is able to understand how her words express certain attitudes of hers, and she is able to correct her linguistic mistakes when pressed to do so by her peers. However, Sandra does not have a clear understanding of many aspects of her new language. The grammatical norms used are a mystery, certain idioms are lost on her, but she is, nonetheless, able to correct a range of mistakes made by her fellow speakers when they perform utterances similar to her past, unsuccessful utterances. This is because she understands and makes use of a variety of locutions, but she is unable to manipulate their individual parts. Suppose, for example, that she can utter and understand a sentence like, "Is the train on time?" but that she has no grasp of which parts of the sentence pick out the concept of "train" or "time." Thus, she can perform various utterances in order to navigate the world around her, but she cannot manipulate the language in the same manner that a fully competent speaker might.

If we were to ask whether Sandra is a fully responsible speaker, my intuition in this case is that she is not. She fails, in my estimation, to meet certain vital conditions for responsibility. However, Sandra also appears to meet all of the criteria for McKenna's notion of responsibility. She is able to understand that others in her linguistic community will interpret her words as expressive of certain mental states, she is able to express her mental states in a variety of ways, and she is able to speak as if her words are bearers of meaning. She might even be able to recognize, in certain contexts, various conversational

implicatures employed by her fellow speakers. However, she is completely in the dark with respect to the structural norms governing her new language. Her mastery of the conventions and cooperative rules of her new linguistic community extends only to the level of functional mastery. She is able to speak *as if* she is a fully competent language user, but in reality she lacks an important standing with respect to her formally trained peers. We are left, then, with a case that is in many ways similar to John Searle's famous "Chinese room argument."[19] Searle's argument was in reference to systems of artificial intelligence, but his point is relevant here as well. What the case of the functional speaker gives us is an agent that can successfully manipulate the signs of a given language yet fails to fully *understand* that language. Searle's claim that such an agent lacks understanding supports my contention that Sandra fails to meet a certain criterion for responsibility. In short, what I am suggesting is this: there are two senses in which a speaker may understand a language. The first is a complete understanding in which the speaker can fully manipulate the various elemental portions of the language in order to employ a nearly endless variety of locutions. The second is a functional understanding in which a speaker has a limited (but perhaps still large) number of locutions at her disposal but lacks the capacity to manipulate the language to its fullest extent (Imagine, here, a speaker who has at her disposal a list of sentences that she understands along with a list of their respective appropriate contexts of utterance, but who is unable to exercise the linguistic dexterity required to use their constituent parts in unique ways. She may be able to use these sentences in order to communicate with her fellow speakers, but she clearly lacks an important capacity as a speaker.) My contention is that McKenna's view requires only the latter communicative

---

[19] Searle, "Minds, Brains, and Programs."

capacity whereas the relevant capacity for a conversational theory of responsible agency is the former.

The above case strikes me as a plausible case of someone who is able to engage in a robust conversation while nevertheless being a less than fully competent (and, thus, less than fully responsible) speaker. A case can be made, I think, for the possibility that there may be similarly 'functional' moral agents. Consider:

> **The Functional Moral Agent:** Imagine that Martin is an agent who lives in a world full of moral norms, but he is unable to recognize the content of those norms. Rather, he observes the actions of others, and understands the ways in which moral agents react to other agents in the world. He realizes that others react to him in predictable ways, and he understands both that his actions are taken by others to communicate certain attitudes and that these interpretations tend to reflect the attitudes he actually holds. Moreover, since Martin has been held responsible by others in the past, he is aware of certain action types that are such that they should elicit certain responses. Thus, he is able to hold others responsible for a variety of actions by reacting to them with appropriate attitudes. Despite all of this, morality, for Martin, is essentially a matter of choreography. He knows what to do or say in response to external stimuli, but he is blind to many of the underlying considerations that a typical moral agent would find compelling.

When we inquire as to Martin's status as a responsible agent, I think that, like Sandra, we will conclude that he is not fully responsible. Martin has all of the abilities necessary to engage in the sort of exchange that McKenna envisions, however something important seems to be missing that, I think, precludes us from counting him as fully responsible. If this is true, then a case like Martin's would surely be problematic for a view like McKenna's as it would show that an agent may have the functional ability to carry on a moral conversation but nevertheless fail to be responsible. There are surely objections that might be made at this point, but I will wait to address these until Section 5, below. For now, I simply want to set the stage for the challenge that ASD presents for the conversational view. While Martin's case is one imagined by a philosopher, it seems to me that the

empirical evidence regarding ASD supports the possibility that there are actual agents in the world who might be considered functional agents in the same way as (though perhaps for different reasons than) my fictional case above. The next two sections will be aimed at showing this to be true.

Section 4: The Moral Psychology of Autism Spectrum Disorder

To show how it might be that autistic individuals could be considered functional agents in the sense described above, I will proceed by highlighting several important features of autistic moral psychology. By attending to the empirical work on the disorder, we will be able to see why autistic individuals might prove to be problematic for the conversational theory.

To begin, autism is generally defined as, "A pervasive developmental disorder characterized by a pattern of deficits that include impaired (delayed and deviant) communication skills; failure to develop social relationships; and restricted, repetitive, and stereotypical behaviors."[20] However, in recent years, a wealth of research has appeared investigating the nature of the disorder and has enabled us to identify the particular sorts of deficits that are present in autistic individuals. Perhaps unsurprisingly, many of them are such that they could drastically alter the way that we think about responsibility. It is important to keep in mind that autism is a spectrum disorder, and, as such, the presence of symptoms can vary widely between individuals. Nevertheless, there are several features that seem to be present in most instances of the disorder. I will begin by paying attention to some features of ASD that are primarily affective in nature. Before doing so, however, a short aside is needed.

---

[20] Accardo & Whitman, *Dictionary of Developmental Disabilities*, pg. 39

The discussion of ASD in this paper will be focused on "high-functioning" autistic individuals. However, there is a great deal of controversy over what, if anything, this phrase means. In the empirical literature, "high-functioning" simply refers to the fact than an individual has an IQ in the normal range or higher. However, there is an important concern that standard tests of IQ require certain interpersonal abilities with which autistic persons struggle, and, so, that these tests stack the deck against autistic persons.[21] Another way that we might understand the phrase "high-functioning" is in relation to an individual's ability to *function* well in the world. That is, we might think that individuals are high-functioning in virtue of the fact that they are able to manage their lives independently and to successfully navigate everyday situations.[22]

In my view, something like this latter understanding is most appropriate, though one likely requires some minimally sufficient level of intelligence to be high-functioning in this sense. For the purposes of my arguments here, however, I will use the "high-functioning" designation simply as a way of picking out those autistic individuals who possess a level of cognitive and executive capacity necessary to engage in moral behavior in the sense to be specified below. I doubt that there is any hard and fast rule that could adequately capture the difference between high- and low-functioning autistic persons in all cases (and I certainly do not wish to take a stand on the issue of whether, or how, the distinction should play a role in policies regarding the treatment or accommodation of autism). I think that it is relatively uncontroversial to suggest that there are autistic persons who are capable of functioning well in the world and who have sufficient cognitive

---

[21] See, for example, Lim, "Accommodating Autistics…"
[22] See Jaarsma and Welin, "Autism as a Natural Human Variation…" and Ibid. for problems with this characterization.

capabilities to engage in (what looks like) moral behavior. My aim is to understand what these individuals can teach us about moral responsibility in virtue of their being autistic rather than in virtue of some other feature of their psychology in addition to autism. My hope is that readers will grant that this may be possible in spite of the problems with the high- vs. low-functioning distinction.[23]

4.1: Affective Features: Empathy, Emotion Recognition, and Hypersensitivity

Although autistic individuals differ in many ways from typically developed people, they have been found to perform equally as well as neurotypical individuals on tests measuring the capacity for emotional empathy. On at least one account, emotional empathy consists in the capacity to experience an emotional reaction in response to the affective state of another,[24] or, as one author puts it, emotional empathy "involve[s] the set of feelings elicited in response to the affective state of others which can carry feelings of warmth or concern *for* an agent (empathic concern) or a set of self-oriented feelings generated *by* such an agent (personal distress)."[25] This is an important point to keep in mind and one that tends to be overlooked in an important way in the philosophical literature.[26]

A 1999 study conducted by R. James Blair[27] provides compelling evidence of this capacity. In the study, Blair tested the skin conductivity responsiveness of several autistic

---

[23] I am grateful to an anonymous referee for pressing me to clarify this.

[24] It should be noted that this conception of emotional empathy may be contentious. Some may argue that a stronger conception is required in which an agent has not only the capacity to have emotional responses but the further capacity to take up another person's emotional perspective. For my purposes in this paper, it will be enough if it can be seen that autistic individuals are emotionally empathic in the weaker sense that I note here. My thanks to David Shoemaker for raising this point.

[25] Gleichgerrcht et. al., "Selective Impairment of Cognitive Empathy."

[26] See, for example, Kennett, "Autism, Empathy, and Moral Agency," in which she treats autistic individuals as lacking empathy, full stop.

[27] Blair, "Psychophysiological Responsiveness"

individuals to expressions of emotional distress in others. Increased skin conductivity is a sign of increased emotional arousal, and what Blair found was that autistic individuals showed increased responsiveness to emotional distress cues in others relative to emotion neutral cues. In discussing the study, Blair writes, "[T]here was some evidence that at least some of the children with autism were finding the pictures of others in distress as aversive. Thus, two of the children with autism tested placed their hands in front of their eyes when a distress cue picture was presented to them and refused to look at it. When asked, they specifically stated that they did not like these pictures."[28] Further evidence of the capacity may be found in a recent study by Dziobek et. al. in which a comprehensive empathy test was administered to individuals with Asperger Syndrome.[29] The results of the test showed that individuals with AS reported comparable levels of both emotional arousal and concern for the emotional distress of others to a typically developed control group. So, at a minimum, we can conclude from these studies that autistic individuals are as capable as neurotypical individuals of experiencing emotional responses toward the emotional states of others.

Another important affective feature of ASD is the fact that autistic individuals seem to have difficulty recognizing others' emotions which, in turn, may lead to a difficulty in making judgments of social appropriateness and in engaging in positive social interactions. Many studies suggest that autistic persons have a substantial degree of difficulty relative to neurotypical individuals in identifying inappropriate social interactions,[30] and one explanation for this feature is that autistic individuals have emotion recognition deficits

---

[28] Ibid., pg. 483.
[29] Dziobek et al., "Dissociation of Cognitive and Emotional Empathy"
[30] See Loveland et. al. (2001); Shulman et. al. (2012).

that make such situations difficult to process. However, there is a substantial degree of disagreement over whether or not such a deficit is actually present in ASD. [31] One possible source of this widespread disagreement may be that studies testing emotion recognition have failed to account for another condition that autistic individuals commonly posses: alexithymia, or, the inability to identify one's own emotions. One recent study[32] claims that what may predict the inability to recognize emotions may be alexithymia and not autism. Moreover, since "studies suggest severe degrees of alexithymia in at least %50 of individuals with autism,"[33] this may explain why there is such a discrepancy in tests of emotion recognition in autism. Whatever its cause, several studies suggest that a deficit in emotion recognition may be a key part of the difficulty that autistic individuals experience with respect to engaging in appropriate (whatever that may entail) social interactions. Nevertheless, recent research suggests that a functional ability may be gained in this respect.

Despite the difficulty in emotion recognition that autistic individuals experience, there has been some promising research recently which suggests that such recognition can be taught through the use of various computer programs.[34] Programs designed to teach facial recognition have been shown to improve the ability of autistic individuals to recognize emotional cues and, as a result, to have more successful social interaction. So, while autistic individuals may have difficulty recognizing and identifying emotions in others, there is good evidence that they can cultivate other heuristic systems in order to allow for more felicitous social interaction.

---

[31] See Uljarevic & Hamilton (2013) for an overview of studies on this point.
[32] Cook et al., "Alexithymia, Not Autism ..."
[33] Ibid., pg. 724.
[34] See, for example, Golan et. al. (2010); Hopkins et. al. (2011).

Learned social interaction of this kind is not necessitated purely by emotion recognition deficits, however. Many autistic persons also suffer, to varying degrees, from emotional hypersensitivity. Emotional hypersensitivity can occur in a number of ways,[35] but one important way occurs when an individual experiences an overwhelming rush of emotions which occur too quickly, or in too great a volume, to be processed adequately. This sort of hypersensitivity is closely related to the alexithymia that many autistic individuals experience, and it leads to a number of important behavioral features. Individuals with emotion hypersensitivity may be more likely to withdraw from social contact or may have difficulty relating to other people as a result of the overwhelming, and often painful, rush of emotional responses. Individuals who are high-functioning, however, are often able to overcome this by developing a number of compensatory strategies (I will say more about these below) in order to allow themselves to interact with their environment.

There is surely more to be said about the emotional capacities and deficits that characterize ASD, but the above will suffice for the purposes of this paper. Before seeing how these features may pose problems for McKenna's theory of responsibility, it will be necessary to say something about the cognitive features of ASD that may be relevant as well. It is to this that I now turn.

4.2: Cognitive Features: Theory of Mind, Moral Judgment, and Compensation

Perhaps the most notable cognitive impairment that is present in autistic individuals is the deficit in cognitive empathy that is consistently observed. These individuals, studies have found, tend to demonstrate impairments in cognitive empathy, or

---

[35] For a compelling, first-personal description of these, see Williams, *Autism: An Inside-Out Approach*.

theory of mind (ToM). This means that autistic persons are unable to represent mental states in other people. That is, they cannot infer the contents of another person's mind. The most famous study in support of this fact comes from Baron-Cohen et. al. In their paper, "Does the Autistic Child have a 'Theory of Mind,'" they developed a task in which participants are asked to answer a question which requires knowledge of another person's false belief. The study found that autistic individuals performed significantly worse than control groups at successfully performing the false belief task, and they inferred from this that autistic participants have significant difficulties in recognizing the contents of other minds, or mind-reading. Subsequent studies have attempted to replicate these findings with a great deal of success.[36]

Another important feature of the moral psychology of autistic individuals is that they are able to recognize at least one important distinction when making moral judgments yet unable to recognize others. Most importantly, autistic individuals tend to be as adept at making the moral/conventional distinction as their typically developing counterparts.[37] That is, they are able to distinguish between transgressions on the part of others which involve what we commonly view as moral wrongs and those which involve wrongs that simply defy convention. It is worth noting that there are a number of difficulties that must be sorted out when it comes to articulating precisely what is involved in this distinction,[38] but I will remain noncommittal on how to address those here. I bring up the moral/conventional distinction simply for the purpose of showing that autistic individuals

---

[36] See, for example , Moran et. al., "Impaired Theory of Mind."
[37] See Blair, "Morality in the Autistic Child"; Leslie et. al., "Transgressors, Victims, and Crybabies."
[38] See, for example, Shoemaker, "Psychopathy, Responsibility, and the Moral/Conventional Distinction."

seem to be sensitive to the same sorts of considerations as the rest of us on this point.[39] Moreover, autistic individuals can recognize the moral/conventional distinction even in cases where authoritative permission is given for the transgression.[40]

Despite being able to recognize the moral/conventional distinction, autistic individuals tend to have problems with other fine-grained features of moral judgment. For example, studies have found that these individuals have trouble distinguishing intentional from unintentional harms.[41] Moreover, they tend to have similar problems distinguishing attempted harms from neutral actions. In other words, when asked to make a moral judgment involving an act in which an agent accidentally causes some harm, autistic individuals will tend to make the same judgment and prescribe the same degree of punishment as they would in cases where an agent intentionally inflicts the same harm. Likewise, when asked to make a moral judgment in cases in which an agent tries but fails to cause harm to another, autistic individuals will tend to make the same judgment as they would in morally neutral cases. This is perhaps not surprising given the deficit in cognitive empathy experienced by these individuals. In cases of accidental harms or failed attempts to intentionally harm, neurotypical individuals tend to assign blame or punishment for what is seen as a morally objectionable attitude or quality of will. However, determining the contents of an attitude or the quality of one's will requires the ability to infer some fact about the mental state of the person in question, and, as we have seen, this is an ability that many autistic individuals seem to lack.

---

[39] Another way of interpreting this evidence is simply to conclude that these individuals are in tune with what matters with respect to morality. It may be the case, however, that the ability to understand the content of moral norms is not the relevant capacity for morally responsible agency. I will return to this issue below. For now, the important point is that autistic individuals are at least sensitive to moral constraints in this, perhaps limited, way. I'm grateful to Michael McKenna for pressing this objection.

[40] Zalla et al., "Moral Judgment in Adults with Autism Spectrum Disorder."

[41] See Moran, et. al. (2011); Koster-Hale et. al. (2013).

Importantly, in most every study on moral judgment in high-functioning autistic individuals the results indicate that these individuals struggle to articulate adequate justifications for their moral judgments. In many cases, the explanation for any given moral judgment relies on an appeal to some rule or another and this in a very general condemnatory manner (e.g. "It's bad to lie."). One way of explaining this might be to appeal to the fact that high-functioning autistic individuals become high-functioning in virtue of their ability to develop compensatory systems that make up for some of the cognitive or emotional deficiencies that I have outlined here. In order to make sense of their surroundings and to function in the world around them, autistic individuals must come up with a way to relate to the world in spite of the psychological factors associated with ASD. One way that they are able to do this is by identifying a set of rules by which they can orient their interactions with other people. As a result, autistic individuals are often seen as rigid rule-followers,[42] and this would explain the self-reported justifications for moral judgments like those just mentioned. More generally, it is important to recognize that individuals on the high-functioning end of the autism spectrum cope with their autism and with their surroundings by developing cognitive strategies which allow them to circumvent some of the difficulties that their disorder entails.

Section 5: The Challenge for the Conversational Theory

With these features of autistic psychology in mind, I would like to now make explicit the challenge that they pose for a view like McKenna's. In other words, I am now in a position to say just how an autistic agent might be similar to the functional agent described in Section 3. Given the fact that there may be autistic individuals who exhibit a combination

---

[42] For more on this, see McGeer, "The Varieties of Moral Agency."

of the symptoms outlined above, it seems to me that such individuals meet all of the criteria for moral responsibility that the conversational view necessitates (i.e. they possess all of the relevant capacities that are needed to make one morally responsible) yet who are nevertheless not responsible in important ways. If this is true, then it may well be the case that there are actual individuals in the world who, upon investigation, show an important gap in McKenna's view. To see how this may be, consider the following hypothetical case:

> **The High-Functioning Autistic Agent**: Adam is an individual on the high-functioning end of the autism spectrum. His autism causes him to experience emotional hypersensitivity, and he has often been overwhelmed with what he experiences as unstoppable waves of emotional responses. In order to allow himself to function well in the world, Adam has developed a compensatory system which allows him to make sense of his emotions by adhering strictly to a set of rules that he has identified over time. He often has emotional responses to other people when they show signs of distress, but he is alexithymic and has difficulty identifying the emotions he experiences in response to others as well as in recognizing the emotions of other people. However, Adam has had lots of practice in social settings and is able to draw on his knowledge of prior experiences in order to engage in 'appropriate' social interactions. Further, let us stipulate that Adam is capable of making reliable moral judgments but that he has a deficit in cognitive empathy as is characteristic of ASD.

Now, according to McKenna's conversational theory, the moral responsibility exchange is underpinned by both the modest metaphysical interpretation of the relationship between being responsible and holding responsible as well as a broadly Strawsonian quality of will thesis. This being the case, I would now like to examine whether or not someone like Adam can rightly be said to engage in the exchange that McKenna describes.

To begin, consider the passage from McKenna quoted above. He begins by saying, "when a morally responsible agent acts, she understands the standpoint of holding morally responsible..."[43] Can this understanding be attributed to Adam? It certainly seems so. From the fact that Adam is able to make reliable moral judgments and distinguish them from

---

[43] McKenna, *Conversation and Responsibility,* 99.

judgments of convention, it seems to follow that he is able to recognize when others have violated moral norms. Moreover, what seems to be suggested by the empirical literature on autism is that the capacity to make reliable moral judgments stems from the capacity to experience emotional empathy, a capacity which, I have stipulated, Adam possesses. If this is true, then it would suggest that Adam is, as Russell requires, able to "feel and understand moral sentiments or reactive attitudes." That is, his ability to have an affective response to others suggests that Adam is in fact capable of experiencing the reactive attitudes.[44] Now, it might be objected at this point, that simply experiencing an affective response is not sufficient for holding morally responsible and that instead, the agent must be feeling a particular reactive emotion. This strikes me as entirely true, and I will return to this objection below.

In addition to being able to hold others responsible, McKenna claims that, "A morally responsible agent must be able to *act as if* her actions were morally significant bearers of meaning among communicating agents."[45] (Recall also, that agent meaning, on McKenna's account, involves reference to the quality of the agent's will.) Can this ability be attributed to someone like Adam? It might seem that the answer is "no," and that this is where McKenna's theory is able to account for autistic individuals. The reason for this is simple. Being able to express meaning, it might be argued, requires the ability to understand how others will interpret one's actions, which requires the ability to represent the mental states of others, and this is precisely the ability that autistic individuals lack.

---

[44] I've left it an open question whether or not his experience of these emotions is in any way standard or whether he experiences reactive emotions in a way that is phenomenologically like the way in which neurotypical individuals experience them. However, on a view like McKenna's which is concerned with a primarily functional capacity (as I argue his is), it may not be important that autistic individuals experience the same emotions as the rest of us, so long as they understand the behavioral manifestations that their emotions call for in a given context. I return to this below.

[45] Ibid. Emphasis added.

However, I want to resist this characterization. More specifically, I want to deny that an understanding of another individual's mental state is necessary for one to act as if one's actions are bearers of meaning.

As was noted above, recent research suggests that autistic individuals are capable, to a certain degree, of learning how to engage in felicitous social interaction. That is, they are capable of recognizing certain features of others which will allow them to successfully engage in social transactions, and I have ascribed this ability to our fictional agent, Adam, as well. This, so far as we know, does not constitute learning cognitive empathy. Rather, it seems to be a type of functional learning which eases social facility. I have also pointed to some evidence which suggests that these individuals are capable, either consciously or unconsciously, of developing compensatory methods that lead to a greater degree of social awareness and appropriate social interaction. The upshot of these facts for the present discussion is that such individuals seem to be perfect examples of agents who can perform meaningful actions but who nevertheless lack an important capacity that is directly related to responsibility. This is partially because the compensatory system that someone like Adam develops is a purely functional system. That is, it is a system which allows him to engage successfully and prudentially with the outside world, but it is not such that it negates the incapacities associated with ASD, nor is it such that it can succeed in all morally relevant situations. If it were, then we might be tempted to say that autistic individuals who have developed compensatory strategies are fully responsible. However, this strikes me as the wrong conclusion to draw for reasons that I will make clear presently.

Astute readers will notice a tension in what I have so far written. I have claimed that the ability of autistic individuals to make the moral/conventional distinction along with

their capacity to have emotional responses to the affective states of others suggests that these individuals are able to hold others responsible. However, I have also just claimed that the compensatory heuristics developed by high-functioning autistic individuals are such that they allow for the possibility of engaging in a moral conversation like the one McKenna has in mind. If the former is true, then it may imply that the actions of autistic individuals do not actually implicate their quality of will because the conversational theorist may simply argue that having an emotional response and recognizing moral transgressions do not constitute understanding the standpoint of holding responsible (this was the problem that I noted in the text above as well as in note 35). On the other hand, if the latter claim is true, then it may imply that autistic individuals are indeed responsible and that their autism is simply a mitigating factor in their responsibility.[46] In short, it may be that I am either describing a class of agents who lack precisely those capacities that McKenna identifies or that I am describing a class of people who are simply responsible agents whose disorder may, in some cases, provide them an excuse. That being said, I take myself to be doing neither of these. Rather, what I am after is a description of an agent who occupies a middle ground. That is, I am looking for a case of someone who is functional in such a way that her quality of will is implicated in her actions but lacks, at least intuitively vital capacities for responsibility – someone who is *merely* functional in this way.

To answer these objections, we must consider how McKenna proposes that we understand the nature of emotions – by first understanding their fitting behavioral manifestations. He writes, "I want to resist [the] picture of the emotions … in which private episodes and subjective states are more fundamental. In my estimation … the order of

---

[46] I'm grateful to both David Shoemaker and Michael McKenna for raising this worry.

conceptual priority is reversed … What makes a private episode of, say, resentment intelligible to its subject *as* resentment is precisely an appreciation of the criteria indicators that would be manifested in a public display of *that* emotion."[47] To see why someone like Adam might prove difficult for McKenna's view to explain in this regard, consider the response that Adam might have to being held responsible by a peer. I have stipulated that Adam is alexithymic, so let us suppose that upon being held to account he experiences a strong emotional response (as a result of his capacity for emotional empathy), but he simply cannot identify the emotion that he feels. However, Adam has been able to compensate for his emotional experiences by tracking past outward manifestations of them, and these outward manifestations are precisely what McKenna sees as important, at least insofar as they are typically associated with the relevant internal states. My contention, then, is that it is plausible to hold that someone with Adam's moral psychology could make sense of the appropriate social practices governing a situation and not only understand but also *identify* the emotion that he is experiencing in light of an understanding of the actions that are called for in a given situation and that these could lead him to identify the internal state accordingly (imagine this as a sort of emotional inverse spectrum problem). Thus, for Adam, even though he may be unable to identify the emotional response that he has in a given situation, his functional capability to recognize socially appropriate transactions would allow him to do so and would thus entail that he *actually does feel the relevant emotion*. If this is true, and I think that it must be on McKenna's view, then we have an agent who meets all of the requirements for fully responsible agency set out by the conversational theory but who seems not to be fully

---

[47] Ibid. pg. 69.

responsible, and, therefore, an agent who occupies the middle ground that I proposed above. He lacks access to a range of facts about other people which may be necessary to meet certain epistemic conditions of responsibility, and more importantly, he lacks the ability to relate to his own emotions in important ways. He can have appropriate emotional responses, but he simply cannot make sense of them in the way required for being fully morally responsible. Instead, Adam is an agent who has developed a way to respond to the world around him in ways that allow him to make sense of his environment and of his relationships with others. However, there is a sense in which his actions are a part of what he takes to be a grand moral choreography. He is functional enough to take part in the choreographed moves, and carrying on moral conversations is an important part of this. However, the ability to function in these conversations fails to explain important factors outside of this functionality, and these factors are such that they make a difference for how we understand responsible agency.

6. Concluding Remarks

To summarize, it seems to me that it is plausible to believe that there high-functioning autistic persons who can provide us with good reasons for thinking that the conversational theory of responsibility is mistaken. Individuals such as my hypothetical agent, Adam, are, I think, perfectly able to engage in the type of conversational exchange that McKenna suggests is necessary for morally responsible agency. However, the conversation in which these individuals engage is lacking something important. It is similar to the linguistic conversations that an individual who has been transplanted from her native country to a region where she does not speak the indigenous language might engage in. Such an individual might, after time, learn enough of the language to function in her new

environment and yet lack an understanding of the norms governing the language and those who speak it. Hers would be a functional knowledge. Agents with high-functioning autism, I have argued, engage in a moral responsibility exchange that is functional in just this way. The moral psychology associated with ASD is such that, I believe, high-functioning autistic individuals are able to reliably make moral judgments, experience moral emotions, and engage in social interactions (albeit with great difficulty and sometimes limited success). Yet autistic individuals are closed off from the rest of the moral community in important ways as a result of their relative deficiency in cognitive empathy as well as the way in which they lack access to their own emotional lives. The fact that the conversational theory must overlook these facts and count autistic individuals as fully responsible is, I think, putative reason to conclude that the theory is incomplete in an important way.

BIBLIOGRAPHY

Accardo, Pasquale J. and Barbara Y. Whitman. *Dictionary of Developmental Disabilities Terminology, 2nd Edition*. Baltimore, MD: Paul H. Brookes Publishing Co., 2002.

Baron-Cohen, Simon, Alan M. Leslie, and Uta Frith. "Does the autistic child have a 'theory of mind'?" *Cognition*. Vol. 21. (1985): 37-46.

Blair, R. James R. "Brief Report: Morality in the Autistic Child." *Journal of Autism and Developmental Disorders.* Vol. 26, No. 5. (1996): 571-579.

Blair, R. J. R. "Psychophysiological responsiveness to the distress of others in children with autism." *Personality and Individual Differences*. Vol. 26. (1999): 477-485.

Cook, Richard, Rebecca Brewer, Punit Shah, and Geoffrey Bird. "Alexithymia, Not Autism,

Predicts Poor Recognition of Emotional Facial Expressions." *Psychological Science*.

Vol. 24, No. 5. (2013): 723-732.

Darwall, Stephen. The Second-Person Standpoint: Morality, Respect, and Accountability.

Cambridge, Massachusetts: Harvard University Press, 2006.

Dziobek, Isabel, Kimberly Rogers, Stefan Fleck, Markus Bahnemann, Hauke R. Heekeren,

Oliver T. Wolf, and Antonio Convit. "Disociation of Cognitive and Emotional Empathy

in Adults with Asperger Syndrome Using the Multifaceted Empathy Test (MET)."

Journal of Autism and Developmental Disorders. Vol. 38. (2008): 464-473.

Gleichgerrcht, Ezequiel, Teresa Torralva, Alexia Rattazzi, Victoria Marenco, Maria Roca, and

Facundo Manes. "Selective impairment of cognitive empathy for moral judgment in

adults with high functioning autism." *Social, Cognitive, and Affective Neuroscience*.

(2012): 1-9.

Golan, Ofer, Emma Ashwin, Yael Granader, Suzy McClintock, Kate Day, Victoria Leggett, and

Simon Baron-Cohen. "Enhancing Emotion Recognition in Children with Autism

Spectrum Conditions: An Intervention Using Animated Vehicles with Real Emotional

Faces." *Journal of Autism and Developmental Disorders*. Vol. 40. (2010): 269-279.

Hopkins, Ingrid Maria, Michael W. Gower, Trista A. Perez, Dana S. Smith, Franklin R.

Amthor, F. Casey Wimsatt, and Fred J. Biasini. "Avatar Assistant: Improving Social

Skills in Students with an ASD Through a Computer-Based Intervention." *Journal of

Autism and Developmental Disorders*. Vol. 41. (2011): 1543-1555.

Jaarsma, Pier and Stellan Welin. "Autism as a Natural Human Variation: Reflections on the Claims of the Neurodiversity Movement." *Health Care Analysis*. Vol. 20. (2012): 20-30.

Kennett, Jeanette. "Autism, Empathy, and Moral Agency." *The Philosophical Quarterly*. Vol. 52, No. 208. (2002): 340-357.

Koster-Hale, Jorie, Rebecca Saxe, James Dungan, and Liane L. Young. "Decoding moral judgments from neural representations of intentions." *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 110, No. 14. (2013): 5648-5653.

Leslie, Alan M., Ron Mallon, and Jennifer A. DiCorcia. "Transgressors, victims, and cry babies: Is basic moral judgment spared in autism?" *Social Neuroscience*. Vol. 1. (2006): 270-283.

Lim, Chong-Ming. "Accommodating Autistics and Treating Autism: Can We Have Both?" *Bioethics*. Vol. 29, No. 8. (2015): 564-572.

Loveland, Katherine A., Deborah A. Pearson, Belgin Tunali-Kotoski, Juliana Ortegon, and M. Cullen Gibbs. "Judgments of Social Appropriateness by Children and Adolescents with Autism." *Journal of Autism and Developmental Disorders*. Vol. 31, No. 4. (2001): 367-376.

McGeer, Victoria. "The Varieties of Moral Agency: Lessons from Autism (and Psychopathy)." In *Moral Psychology Volume 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development.* 227-257. Edited by Walter Sinnott-Armstrong. Massachussetts: MIT Press, 2008.

McKenna, Michael. *Conversation and Responsibility.* New York: Oxford University Press, 2012.

Moran, Joseph M., Liane L. Young, Rebecca Saxe, Su Mei Lee, Daniel O'Young, Penelope L. Mavros, and John D. Gabrieli. "Impaired theory of mind for moral judgment in high-functioning autism." *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 108, No. 7. (2011): 2688-2692.

Russell, Paul. "Responsibility and the Condition of Moral Sense." *Philosophical Topics.* Vol. 32, Nos. 1 & 2. (2004): 287-305.

Searle, John. "Minds, Brains, and Programs." *The Behavioral and Brain Sciences*. Vol. 3. (1980): 417-457.

Shulman, Cory, Ainat Guberman, Noa Shiling, and Nirit Bauminger. "Moral ad Social Reasinging in Autism Spectrum Disorders." *Journal of Autism and Developmental Disorders*. Vol 42. (2012): 1364-1376.

Shoemaker, David. "Psychopathy, Responsibility, and the Moral/Conventional Distinction." *The Southern Journal of Philosophy*. Vol. 49, Spindel Supplement. (2011): 99-124.

Strawson, P. F. "Freedom and Resentment." 1962. In *Free Will, 2nd Ed.,* 72-93. Edited by Gary Watson. New York: Oxford University Press, 2003.

Uljarevic, Mirko and Antonia Hamilton. "Recognition of Emotions in Autism: A Formal Meta-Analysis." *Journal of Autism and Developmental Disorders*. Vol. 43. (2013): 1517-1526.

Watson, Gary. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In *Agency and Answerability: Selected Essays.* New York: Oxford University Press, 2004.

Williams, Donna. *Autism: An Inside-Out Approach*. London: Jessica Kingsley Publishers Ltd., 1996.

Zalla, Tiziana, Luca Barlassina, Marine Buon, and Marion Leboyer. "Moral judgment in adults with autism spectrum disorders." *Cognition*. Vol. 121. (2011): 115-126.

Zimmerman, Michael J. *An Essay on Moral Responsibility*. Totowa, New Jersey: Rowan and Littlefield, 1988.