

Responsibility and the Condition of Moral Sense

Paul Russell

University of British Columbia

O wad some Pow'r the giftie gie us
To see oursels as ithers see us!

—Robert Burns

Recent work in contemporary compatibilist theory displays considerable sophistication and subtlety when compared with the earlier theories of classical compatibilism. Two distinct lines of thought have proved especially influential and illuminating. The first developed around the general hypothesis that moral sentiments or reactive attitudes are fundamental for understanding the nature and conditions of moral responsibility. The other important development is found in recent compatibilist accounts of rational self-control or reason responsiveness. Strictly speaking, these two lines of thought have developed independent of each other. However, in the past decade or so they have been fused together in several prominent statements of compatibilist theory. I will refer to theories that combine these two elements in this way as RS (Reason-Sentiment) theories. RS theories face a number of familiar difficulties that relate to each of their two components. Beyond this, they also face a distinct set of problems concerning how these two main components relate or should be *integrated*. My concerns in this paper focus primarily on this set of problems.

According to one version of RS compatibilism, the role of moral sentiments is limited to explaining what is required for *holding* an agent responsible. In contrast with this, the role of reason responsiveness is to explain what moral capacities are required for an agent *to be responsible*, one who is a legitimate or fair target of our moral sentiments. More specifically, according to this view, moral sense is *not* required for rational self-control or reason responsiveness. There is, therefore, no requirement that the responsible agent has some capacity to feel moral sentiment. Contrary to this view, I argue that a responsible agent must be capable of *holding* herself and others responsible. Failing this, an agent's powers of rational self-control will be both limited and impaired. Insofar as holding responsible requires moral sense, it follows that *being* responsible also requires moral sense. Moral sense is, therefore, a condition of responsible agency.

I.

The most influential contemporary statement of the moral sentiment approach to issues of free will and moral responsibility is presented in P. F. Strawson's seminal paper "Freedom and Resentment."¹ The key idea in this paper is that the nature and conditions of moral responsibility must be understood in terms of our reactive attitudes or moral sentiments. The attitudes and intentions of our fellow human beings, Strawson maintains, naturally and inescapably arouse reactive attitudes and feelings in us. Incompatibilist concerns about the implications of determinism can neither dislodge nor discredit our commitment to feelings and attitudes of this kind. Although our particular moral sentiments require some relevant justification, and issues of excuse and mitigation arise with respect to them, there is no general, external justification required for the whole framework of moral sentiments. This is a given of our human nature. These observations about the natural operation of the human mind serve as a barrier to all systematic skeptical worries about the foundations of moral life.² These foundations are not based on a general rationale of some kind (e.g., a hypothesis about contra-causal freedom, etc.) but rather depend on the natural workings of human feeling and emotion.

Strawson's approach distinguishes between two ways in which agents may be excused for their actions. The first are specific excuses, such as ignorance and accidents, and the second are exemptions, where the agent is judged to be morally incompetent. In the case of specific excusing considerations there is no suggestion that the agent is in any way an inappropriate object of moral sentiment. All that is claimed is that the particular conduct in question does not manifest any lack of due care or regard on the agent's

part.³ In contrast with this, when exempting considerations are applicable we are invited “to suspend our ordinary reactive attitudes toward the agent” on the ground that they are in some relevant respect abnormal or immature.⁴ The account of exemptions is crucial to Strawson’s approach because it serves to identify the *boundary* between responsible and nonresponsible agents. That is, exemptions identify those individuals who are not moral agents or full members of the moral community as delineated by the legitimate scope of our moral sentiments. It is here, however, that critics have argued that Strawson’s theory runs into fundamental difficulties.

Clearly any theory of the kind that Strawson is advancing must provide some plausible interpretation of moral capacity of a kind that allows us to make sense of exempting conditions. This is something that Strawson fails to provide, which opens the door for the incompatibilist to argue that the relevant capacity is some mode of libertarian free will or contra-causal freedom (this being what the “immature” and “abnormal” lack). As it stands, Strawson has no convincing line of reply to this criticism. What is needed, therefore, if Strawson’s approach is to succeed, is some relevant compatibilist account of moral capacity.⁵

A related way of describing this general objection to Strawson’s approach is to argue that his approach involves a fundamental confusion.

Strawson’s theory may reasonably be said to give an account of what it is for agents to be held responsible, but there seems to be a difference between being *held* responsible and actually *being* responsible. Surely it is possible that one can be held responsible even though one in fact is not responsible, and conversely that one can be responsible even though one is actually not treated as a responsible agent. By understanding responsibility primarily in terms of our actual practices of adopting or not adopting certain attitudes towards agents, Strawson’s theory risks blurring the difference between these two issues.⁶

The exact nature of this criticism needs careful interpretation. It is clear that Strawson wants to allow that in particular cases our reactive attitudes may fail to track moral responsibility (e.g., we feel a sentiment of blame because we have incorrect beliefs about the agent’s intentions or state of mind). This is not where the objection lies. The force of the objection is that our established attitudes and practices may be *systematically* misplaced or misdirected. For example, the fact that a moral community actually *holds* juveniles or mentally incompetent individuals as responsible does not show that these individuals *are* responsible. The way in which our moral sentiments are directed in cases of this kind cannot, in itself, serve as a reliable guide to a standard of responsibility that is reflectively legitimate. What we need, therefore, is some independent account of moral capacity and the conditions of moral agency that can guide our moral sentiments.⁷

II.

According to prominent versions of RS theory, these weaknesses in Strawson's approach can be eliminated by reason responsive or rational self-control theories, which serve as the basis of a more satisfactory account of moral capacity. That is to say, what the moral sentiment approach plainly needs is a more developed theory of the boundary between normal/abnormal and mature/immature as it concerns who is or not an appropriate target of our reactive attitudes. Whereas Strawson's remarks on this subject are too thin and slight to do this job, reason responsiveness seems to provide exactly what we are looking for.⁸ This aspect of RS theory has been worked out in detail in the recent and highly influential theory of John Fischer and Mark Ravizza. What sort of conditions must be satisfied for an agent to be an appropriate or fair target of reactive attitudes? According to Fischer and Ravizza a responsible agent is one who can in some suitable sense *control* his conduct. They refer to this as "the control condition" (RC, 13–14).⁹ What is required is that the responsible agent must be able to both recognize and react to available reasons. In the case of reason recognition it is essential, on this account, that there is some pattern or regularity to the way that the agent recognizes reasons. (For example, it will not suffice that the agent is occasionally or intermittently able to recognize relevant reasons for action.) However, it is not essential that the agent is regularly guided by reasons in this way. Reason reactivity is satisfied, they argue, when the agent's conduct shows that his conduct has on occasion been guided by his reasons.¹⁰

On the face of it, this account of reason responsiveness provides RS theory with a general interpretation of moral capacity of the kind that was missing from Strawson's discussion. The question that now arises, however, is whether there is more required of responsible agency than satisfying reason responsiveness. Fischer and Ravizza argue that there are further and distinct requirements to be met for the agent to be in control of her conduct. More specifically, the agent must not operate with reason responsive dispositions that have themselves been manipulated or implanted in such a way that she may be covertly controlled.¹¹ In other words, we must address concerns about the historical process by which the agent has acquired her reason responsive dispositions. In circumstances where implantation or manipulation of some kind has occurred, the agent cannot be said to "own" the disposition that she is operating with. If illegitimate processes of this kind have occurred, then responsible agency is compromised.

How, then, is "ownership" to be accounted for? Whether our reason responsive dispositions are owned or not will depend on the particular *history* involved in their acquisition. There are, according to this theory, three necessary steps required for ownership to take place. The first is that the agent

(i.e., the early stages of childhood) comes to see herself as an agent who produces “upshots” in the world.¹² When this stage is reached, then the child is in a position to see herself as a “fair target of reactive attitudes as a result of how [she] exercises this agency in certain contexts.”¹³ Taken together, these two steps require that the responsible agent must have a certain view of herself as an agent and as a fair target of the reactive attitudes. Following Galen Strawson, Fischer and Ravizza describe this as committing themselves to a *subjectivist* approach to moral responsibility.¹⁴ There remains the difficulty, however, that these first two steps in the process of “taking responsibility” may themselves be manipulated or covertly controlled. To deal with this problem a third step is necessary. The cluster of beliefs specified in the first two steps “must be based, in an appropriate way, on the individual’s evidence.”¹⁵ What this is intended to exclude is any case where there is direct implantation or manipulation of the agent’s reason responsiveness. In other words, it is essential that the historical process by which an agent acquires their reason responsive dispositions and takes responsibility for them must come about in some *appropriate* way. What is “appropriate” does not exclude causal determinism but does exclude all forms of manipulation, implantation, or process that permits covert control.¹⁶

It is clear, then, that on this account more than reason responsiveness is required for moral agency and responsibility. The responsible agent must be guided by reason responsive dispositions that have been produced by some “appropriate” historical process. The success of this appeal to history and the process of “taking responsibility” must be judged in terms of whether or not we provided a principled basis for distinguishing between “appropriate” and “inappropriate” historical processes.¹⁷ It is, however, at exactly this point that the argument runs dry. Fischer and Ravizza leave the relevant distinction unanalyzed and thereby ask the reader to accept that there is some intuitive distinction between cases whereby the acquisition of reason responsive dispositions is merely determined and cases where it is manipulated or covertly controlled in some way.¹⁸ This is not, however, a distinction that incompatibilists will allow to serve as a boundary between responsible and nonresponsible agency. The reason for this is that what fundamentally troubles incompatibilists about manipulation and implantation cases is not the matter of “ownership” as such, but that agents in these circumstances are operating with reason responsive dispositions that they have acquired through processes that they have *no control over*. From this point of view it is not enough that agents are reason responsive and have gone through an “appropriate” process of “taking responsibility.” The fact remains that—just as in manipulation and implantation cases—even these “normal” agents have no (final or ultimate) control over the kinds of dispositions that guide their conduct. Unless compatibilists are able to show that an “appropriate” history

serves to distinguish between agents who have control over their reason responsive dispositions and those who do not, it follows that appeals to history are irrelevant.¹⁹

III.

In the previous section we reviewed the basic features of RS theory with particular reference to the way that Fischer and Ravizza have recently presented it. The conclusion that we reached was that although RS theories need to address the general problem of manipulation and implantation, appealing to history and the process of “taking responsibility” looks like an unpromising way of dealing with these concerns. For this reason I have expressed skepticism about the rationale that Fischer and Ravizza provide for taking a “historical” approach to RS theory. Having said this, it is important not to lose the baby with the bathwater when considering the importance of the “subjective” requirement that has been proposed for moral agency. The contrast between versions of RS theory offered by Fischer and Ravizza, on one hand, and Wallace, on the other, bring this out. Although Wallace’s theory of rational self-control shares the same general features as the account of reason responsiveness offered by Fischer and Ravizza, Wallace does not commit himself to any further historical requirements relating to “ownership” issues. More specifically, Wallace’s version of RS theory has nothing that corresponds to the “subjective” requirement that is involved in the first two stages of “taking responsibility,” as described by Fischer and Ravizza. All that is required, on Wallace’s account, is that the moral agent is capable of rational self-control. There is no further requirement that the agent is also capable of seeing herself as an agent who is a fair target of reactive attitudes or moral sentiments. While it is true, on Wallace’s account, that an agent can be *held* responsible only when she is a target of these attitudes and sentiments, her *being* responsible is not a function of her possessing any capacity to hold herself responsible in these terms.²⁰ In contrast with this, Fischer and Ravizza are committed to the view that if the agent has no capacity for moral sense then she cannot be said to “own” her reason responsive dispositions, which will compromise the agent’s responsibility for the conduct that flows from these dispositions.

The question that needs to be asked at this juncture is whether or not there is any *independent* rationale for the subjectivist requirement that the agent must be able to see herself as an appropriate target of moral sentiments—one that has nothing to do with “historical” concerns about “ownership.” We may begin to answer this question by way of observing some problems that arise for RS theory when this subjective requirement is not

met. It should be noted, in the first place, that there is a significant difference in the capacities required for holding and being responsible. Wallace makes clear that his understanding of the general capacity for rational self-control is Kantian in character.²¹ In contrast with this, the Strawsonian role of moral sense—which has evident Humean roots—is limited to explaining what is involved in holding agents responsible. It is a striking fact, therefore, that the Strawsonian/Humean element of moral sense plays no role in Wallace's Kantian account of moral agency or rational self-control. This is indicative of another, related, problem in Wallace's RS theory. Because there is no requirement that responsible agents be capable of moral sense, it is entirely possible that we could encounter rational self-controllers who are responsible agents but who nevertheless have no capacity to *hold* themselves or others responsible.²² In other words, on this theory a significant *asymmetry* may open up between those individuals who are responsible agents (legitimate targets of moral sentiment) and those agents who are able to hold themselves and others responsible. Membership in the moral community may divide and fail to overlap in respect of these two aspects of moral capacity. This is a puzzling result and it raises the question if asymmetrical theory of this kind is ever acceptable. Fischer and Ravizza, however, offer no principled reason for rejecting a theory on this basis. Nor do they take the view that reason responsiveness, as such, will be compromised in circumstances where an agent lacks moral sense or any capacity for reactive attitudes. Their reasons for insisting on the importance of the subjective requirement lie elsewhere.

IV.

The proposal I want to consider is whether or not a capacity to hold oneself and others responsible is a required condition of being responsible. Stated in these general terms, we may call this a condition of *symmetry*: an agent cannot be responsible unless she is also able to hold herself and others responsible. This condition takes a more specific form in RS theory, since it is committed to the more specific view that moral sense is the basis for holding agents responsible. In the case of RS theory, therefore, the condition of symmetry takes the following form:

The responsible agent must be able to feel and understand moral sentiments or reactive attitudes.

Let us call this *the condition of moral sense*. It is possible to reject this condition without rejecting the condition of symmetry. This is possible, however, only if we do not interpret holding responsible in terms of moral sentiments. Among those who reject the condition of moral sense there are some who

nevertheless accept the suggestion that holding an agent responsible involves reference to moral sense.²³ The specific difficulty of asymmetry will arise for those who are committed to this analysis of holding responsible (which is true of all RS theorists).

Why should we accept the condition of moral sense (i.e., if we do not do so on the basis of ownership and historical considerations)? It has recently been argued by Ishtiyaque Haji that an agent can “lack the capacity to hold oneself or others responsible without this lack in any way impinging freedom, epistemic, or authenticity requirements of responsibility.”²⁴ The general assumption operating here is that moral sense is in no way required for rational self-control or reason responsiveness, as these are distinct issues. Contrary to this view, I will argue that there is a more intimate connection between moral sense and the kind of rational self-control that enables us to grasp and be guided by moral reasons. Agents who lack moral sense are missing something that is vitally important and that “normal” moral agents do and must 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.

To help us appreciate this point, consider the parallels between fear and reactive attitudes.²⁵ Take the case of Jill, who is incapable of feeling or experiencing fear. It is not simply that she can control her fear, or “overcome” this emotion, it is that she cannot feel it in any circumstances or conditions. There is no question of her controlling or overcoming this emotion because it *never* arises in her. Clearly Jill is, in this respect, emotionally abnormal. Jill is, nevertheless, a highly intelligent person who can identify what conditions are *dangerous* and may *harm* her. Her intellectual abilities are such that she can describe objects and circumstances of this kind and discourse about them. Jill can also anticipate how normal human beings will react to these circumstances and objects, although she does not know what it is like for them to feel or experience fear or be moved by it. In this sense, Jill has an entirely “external” or “superficial” understanding of fear. Whereas fear, as we might say, *colors* the normal person’s world, and draws attention and interest to dangerous or harmful objects and situation, this is not possible for Jill. Jill’s world is monochromatic in this respect—lacking any emotional highlights to flag certain objects and situations. Since fear plays no role in Jill’s life, she cannot be reached through this emotional route as it is completely closed off to her.

It is evident that Jill is not a full participant in normal human life insofar as we reason and engage with each other about dangerous and harmful situations and objects. Whereas normal human beings live in a world that is colored by emotional tones that give *salience* and *significance* to danger and harm, Jill lives in a black-and-white world that lacks these emotional tones. It would, of course, be a mistake to treat Jill as if she were psychotic or com-

pletely stupid and unable to recognize or converse with us *at all* about these matters. Clearly we can find points of common understanding and shared experience by which Jill can be (hopefully) motivated to avoid danger and harm, relying on channels that do not involve the emotional triggers of fear. It remains true, nevertheless, that we have every reason to recognize the significance of Jill's emotional abnormality and the limits that this places on her ability to recognize and respond to danger.²⁶ Jill's emotional abnormality plainly has considerable *practical* significance for her in situations of this kind.

Consider now the parallel case of Jack, who lacks all capacity for moral sense. Jack is, nevertheless, an intelligent person who *appears* to satisfy reason responsive requirements. That is to say, Jack can recognize and follow moral norms and he is able to anticipate the consequences of any failure to comply with these norms and standards. (Just as Jill is able to identify dangerous situations and avoid them.) The fact remains, however, that violations of these moral norms do not affect Jack in any emotional way. In this sense, Jack is morally *cold*. In these circumstances, an agent such as Jack (as we found in the parallel case with Jill's lack of fear) is not motivated to care about his moral qualities on the basis of his own or others reactive attitudes or moral sentiments. He is unaffected by any consideration of *this kind*. More specifically, it is impossible for Jack to come to "internalize" the reactive attitudes that others may direct at him. It is not just—as may happen with a normal person—that Jack may not *accept* or *agree* with the reactive attitudes that are directed toward him. The problem is deeper than this. Jack cannot even potentially come to share reactive attitudes and feelings because he constitutionally lacks any emotional life of this kind.²⁷ To this extent, therefore, there is no question, for Jack, of accepting or rejecting the reactive attitudes that others direct at him. On this interpretation, it follows that Jack is incapable of any kind of "deep assessment" of himself and others.²⁸

These observations make clear that in an important dimension Jack is not a *full participant* in (normal) moral life. There is a considerable sphere of moral experience and communication that is simply closed off to Jack. In our dealings with any individual of this kind it is both unreasonable and unfair to communicate and reason with him as if this incapacity was irrelevant to this person's ability to function as a moral agent. (Just as it would be unreasonable to treat Jill as if her incapacity for fear was irrelevant to the way she that responds to danger or harm.) At the same time, it is important not to exaggerate this problem (significant as it may be). Clearly Jack is an intelligent person and understands moral rules and expectations and their associated sanctions. For this reason it would also be a mistake to assimilate Jack to other individuals who are wholly unable to participate in the moral community, such as animals or infants. Nevertheless, when it comes to the dimension of moral life that involves reactive attitudes, the most that Jack can do is

“parrot” these responses or feign feelings of this kind. Just as Jill’s world without fear left considerations of danger and harm without any emotional coloring, so too in Jack’s world moral considerations lack any emotional coloring of the kind provided by moral sense. For this reason we must conclude that Jack’s moral world—the way that he experiences it and is moved and directed within it—is *very* different from our own.

It is tempting to present Jack’s lack of moral sense as a completely separate issue from his capacity for rational self-control. The situation, however, is not so simple as this. In the first place, Jack’s moral development must also be fundamentally different from that of the normal person. This will plainly influence the way in which Jack learns to grasp and apply moral considerations. In ordinary moral development there is a close relationship between a child’s evolving ability to understand moral norms and the child’s ability to understand and internalize reactive attitudes when these norms have been violated.²⁹ Clearly Jack’s moral development cannot evolve in this way and this must both limit and alter the way in which Jack eventually becomes able to grasp and be guided by moral considerations.

Apart from developmental issues (e.g., as they relate to moral education), it is important to recognize that Jack’s “control” system, as it relates to moral considerations, is radically different from the normal person’s. On the account given, Jack is able to guide and motivate his conduct with a view to “external” sanctions such as rewards and punishments. Related to this, Jack also has an “external” interest in not arousing negative moral sentiments, since this will obviously affect the way that others treat him. What Jack will lack, nevertheless, is an “internal” system of sanctions (or incentives) as associated with moral sentiments. The internal system operates in a quite different way than its “external” (rewards and punishments) counterpart. More specifically, our experience of negative reactive attitudes is not simply a matter of wanting to avoid the unpleasant or painful consequences associated with them. In cases where we *accept* the response (e.g., blame) as appropriate or fair, something much more complex is going on. In these cases, the agent finds the responses particularly troubling *because she accepts or endorses the moral considerations and norms* that serve as the basis of the adverse reactive attitudes. What this shows is that, for a normal moral agent, our capacity to experience and feel moral sentiments, toward ourselves and others, is intimately and inextricably connected with our understanding of the significance of the background moral claims and considerations.

To appreciate and understand moral considerations fully is precisely to be able to apply them to oneself and others and feel the appropriate way when violations occur. Failing this the agent just “does not really get it.” They are responding to these claims and considerations in an entirely “external” manner. (Compare Jill’s way of responding to danger and harm.) The agent’s general capacity for rational self-control is, therefore, expressed and manifest

through moral sense. Moral sense serves as a “feedback-loop” through which the agent is confronted with the *salience* and *significance* of the moral considerations that she is presented with. Since Jack lacks all this psychological apparatus, it is very evident that Jack’s capacity for rational self-control, in relation to moral life, is significantly impaired and compromised. What this general observation suggests is that it is possible to “over-intellectualize” not just what is involved in *holding* an agent responsible, but also what is involved in an agent *being* responsible (being capable of rational self-control as it relates to moral norms and expectations). Rational self-control is itself dependent upon and *integrated with* our capacity for moral sense.

V.

Critics may argue that these observations about the relevance of moral sense for moral (normative) competence do not matter much. After all, as we have already pointed out, in the normal case, reason responsiveness evolves alongside the development of moral sense, as these are not unrelated components of the moral personality. The distinction between the two capacities is, therefore, both empirically and conceptually unfounded. In reply to this line of criticism, the first thing to be said is that several philosophers have explicitly denied that there is any such interdependence between rational self-control and moral sense—either conceptually or empirically—of the kind that I have described.³⁰ The claims of the moral sense condition in this regard are, therefore, far from uncontroversial. Beyond this, the moral sense condition helps us to understand and interpret a particularly problematic set of cases associated with moral agency. The absence of moral sense, as such, does not make an individual a psychopath. Nevertheless, the absence of moral sense is an especially prominent feature of the psychopathic personality and the behavioral problems associated with it. Psychopathic personalities have, in particular, a pronounced lack of shame and remorse and this reflects wider emotional abnormalities.

What is especially puzzling and problematic about the case of the psychopath is precisely that these individuals appear “normal” and “mature” in respect of rational self-control (i.e., unlike animals, infants, psychotics, etc.).³¹ Several prominent proponents of RS theory have argued that, despite the façade of normative competence, the basis for exempting psychopaths still rests with the way in which their powers of rational self-control have been impaired.³² These accounts, however, make no reference to the specific role that an incapacity for *moral sense* plays here. On the account that I have described, the difficulties that the psychopath faces in respect of rational self-control or reason responsiveness should be understood primarily in terms of

the way in which their impaired moral sense limits and distorts their ability to recognize and react to moral reasons (i.e., as compared with a normal person who is capable of moral sense). For this reason, the account that I have provided is consistent with the general view advanced by Antony Duff that “a psychopath, although not intellectually incompetent, is unable properly to understand the ‘nature and quality’ of his acts, since he cannot grasp those emotional and moral aspects which are as much part of them as their empirical features.”³³

The puzzle with psychopaths is not simply that they are reason responsive and lack moral sense. It is that they have the “mask” or façade of moral understanding despite a lack of moral sense. In most cases where moral sense is absent, it is accompanied by an *obvious* lack of moral understanding or normative competence. The psychopath is not obviously impaired in any systematic way that incapacitates them from understanding and following basic rules and anticipating the consequences of not following them.³⁴ Nevertheless, as I have argued, without moral sense, the way in which this agent becomes sensitive to moral considerations and is motivated to comply with them is very different from the normal, adult person. In particular, while the psychopath may well be motivated to avoid the unpleasant consequences of punishment and the negative treatment that may come with the reactive sentiments of others, this motivation is entirely “external” to the moral considerations that ground these attitudes and practices. What this agent cannot do, unlike his normal counterpart with moral sense, is come to view himself as a target of moral reactive attitudes that are not only unpleasant in themselves, but reflect the fact that he has “internalized” or accepted these reactions to his conduct.³⁵ An agent who is capable of internalizing or accepting reactive attitudes is not only capable of experiencing these emotions and the associated unpleasant feelings, he accepts the legitimacy and significance of the considerations that have produced these feelings (i.e., unlike the agent who has a merely “external” attitude to these sentiments or attitudes). It is this process—the interaction between recognizing and being guided by moral considerations and understanding moral sentiments from the “inside”—that an agent without moral sense cannot benefit or draw from.³⁶

The psychopath’s way of dealing with negative reactive attitudes must always be “external” in character. Naturally these individuals will want to avoid punishment and the unwelcome treatment that comes with negative moral sentiments aroused in others. These individuals may also view these sanctions and the negative consequences that result from their conduct as both predictable and an acceptable part of the rules of society. What these individuals cannot do, however, is accept these reactions and sanctions in the *same way* as a normal person can. Whereas the person with moral sense is capable of accepting these responses in a way that involves coming to feel them from the inside or *sharing* these negative sentiments, this is simply not

possible for the psychopath (or a person like Jack). He cannot, therefore, experience the grip and force of moral considerations through the channel of reactive sentiments themselves. Without an ability of this kind an agent's ability to recognize the salience and significance of moral claims, and to be motivated effectively by them, will be *radically* impaired. (The parallels with Jill's *practical* difficulties in respect of her inability to feel fear and the way that this limits and distorts her responses to danger and harm are obvious enough.)

The above discussion of the case of psychopaths shows that the condition of moral sense has concrete application to an important set of problem cases that we encounter in ordinary moral life.³⁷ Nevertheless, my concern in this context is not to provide an analysis or critique of the complex problem of psychopathy.³⁸ The interest of psychopaths, for our purposes, rests with the light that this sheds on what is required of *normal* moral agents who we regard as fully responsible. For reasons that I have explained, these are agents for whom the functioning of moral sense is directly relevant to the degree to which, and the way in which, their capacity for rational self-control effectively operates. Insofar as impaired moral sense limits or distorts their capacity for rational self-control, this incapacity will limit their ability to operate as full "participants" in the moral community. It is both unfair and unreasonable to treat these individuals who are incapacitated in these ways as if they are fully responsible. Similarly, it is crucial, for the purpose of moral education and understanding, that we carefully explore the intimate relationship between moral sense and rational self-control. In particular, we must avoid any false dichotomy between capacities required for *being* responsible (qua moral agency) and capacities required for *holding* agents responsible. The connection between being a moral agent and being able to hold ourselves and others responsible is much closer than this. Responsible agents must be capable of "deep assessment" in order to be able to effectively recognize and be motivated by moral considerations. Without "deep assessment" there is no "deep understanding" or "true appreciation" of moral considerations. This is what the condition of moral sense speaks to and demands.

VI.

My central concern in this paper has been to argue for the condition of moral sense. As we have noted, a number of philosophers have denied the validity of this condition. This includes several prominent compatibilists who endorse RS theory and accept that moral sense is essential for *holding* people responsible (i.e., for explaining the kind of "deep assessment" that is involved in viewing an agent as being responsible). The moral sense condition is also

a condition of symmetry, as interpreted by any account that takes moral sense to be essential to holding an agent responsible. That is to say, symmetry requires that any responsible agent is also able to hold himself and others responsible. On the assumption that holding responsible requires moral sense, it follows that responsible agency also requires moral sense.³⁹

Related to this point, the condition of moral sense imposes a “subjective” requirement on responsible agency. The responsible agent must be able to see herself as producing “upshots” in the world and as being an appropriate target of reactive sentiments. The basis on which I have defended this “subjective” requirement is, however, very different from the account suggested by Fischer and Ravizza. On their account, the importance of moral sense is that it plays a role in historical considerations as they relate to the issue of “ownership” of reason responsive dispositions. I have expressed doubts about the viability of this approach considered as a way of dealing with worries about implantation, manipulation, and covert control. The basis of my defense of the moral sense condition lies with the role that it plays in the agent’s capacity for rational self-control or reason responsiveness. I have argued that when moral sense is impaired or the agent is incapacitated in this respect then this will directly affect the agent’s ability to grasp and be motivated by moral considerations. Whereas the normal agent who is capable of moral sense will experience and understand reactive sentiments in a way that provides salience and significance to moral considerations, and also provides an internal and independent source of motivation (i.e., as distinct from “external” sanctions) for complying with these demands and expectations, the agent who lacks this general capacity will not be able to employ these resources to govern and guide her conduct. In these circumstances, the agent who lacks moral sense is neither fully responsible nor a full participant in the moral community. By way of analogy, I have suggested, agents who are incapacitated in this way are like individuals who lack any capacity to feel fear in face of danger or harm. Their sensitivity to these considerations is radically and severely impaired and impoverished. The general point that emerges from this defense of the condition of moral sense is that it is important that we avoid “over-intellectualizing” what is involved in *being* responsible—not just what is involved in holding agents responsible.⁴⁰

Although my reasons for accepting the condition of moral sense (viewed as a subjective requirement for responsible agency) are different from those provided by Fischer and Ravizza, I share their view that objections relating to manipulation and covert control must be addressed and that what has been said about reason responsive or rational self-control will not suffice to allay these concerns. It is my view, therefore, that any adequate compatibilist RS theory will need to find some independent way of dealing with (historical) worries of this kind.⁴¹ That said, the observations we have made concerning

the relevance of moral sense for moral agency are by no means without significance for the compatibilist position in its RS form. It is a familiar criticism of compatibilism that its account of responsibility is generally “superficial” and lacks “depth.” It is, therefore, a particular aim of compatibilism to provide sufficient “depth” for its account of the conditions of responsible agency—and reason responsive view has played a prominent role in all this. Related to this search for depth, as one prominent representative of the reason responsive approach has pointed out, is the fact that “complexity” matters.⁴² By failing to recognize the important role that moral sense plays in relation to normative competence and rational self-control, compatibilists have failed to supply the necessary degree of both “depth” and “complexity” that is needed here. Put another way, responsibility is not simply a matter of *narrow depth* as conceived in “intellectualist” terms that ignore the role of emotion in moral life. It is also a matter of *width* as supplied by the interaction between reason and emotion as manifest and expressed in our capacity for moral sense.

This emotional dimension of responsible agency provides something more than an “atomistic” conception of rational competence that makes no reference to moral emotions and the social context in which they are acquired. It draws our attention to the fact that moral competence of the kind required for responsible agency develops in a social and emotional matrix that fosters and nourishes the general capacity to recognize and respond to moral considerations. No adequate theory of responsibility can ignore this important dimension of moral agency and the way that it evolves from childhood to adulthood.⁴³

I have discussed and defended the condition of moral sense almost entirely within a compatibilist framework. The reason for this should be clear. Contemporary compatibilist theory, following the lead of P. F. Strawson, has pursued the theme that moral sense is essential for holding agents responsible. However, Strawson’s theory, as we have also noted, provided little detail in the way of an account of the capacities required to be a responsible agent (i.e., what is required to be a “normal adult”). Contemporary compatibilists have tried to plug this gap in Strawson’s theory with an account of rational self-control that does not itself involve any reference to moral sense. I have argued that this is a mistake and that where moral sense is lacking, rational self-control is seriously impaired and compromised. There is, however, no necessary connection between accepting the moral sense condition and being a compatibilist. That is to say, a libertarian incompatibilist might well agree that moral sense is an essential element in the kind of “deep assessment” involved in holding an agent responsible. The libertarian may also agree that the responsible agent must be reason responsive but will go on to insist that this capacity must include an ability to choose to act on his reasons in a sense that

requires (categorically) open alternatives. Be this as it may, there is nothing about these specific commitments, or the argument that I have presented for the moral sense condition, that does not also hold for the libertarian. That is to say, the responsible agent, on *any* interpretation, must be able to weigh and review moral considerations in light of a general capacity to hold himself and others responsible. Without this capacity this agent will have a *shallow understanding* of the moral considerations that are at stake and will fail to see the salience of these considerations or be motivated by them in the same way as a normal agent (whether she possesses libertarian free will or not). Insofar as an ability to hold oneself responsible requires moral sense, it follows that even a libertarian agent will be significantly incapacitated without moral sense. From this we may conclude that although the condition of moral sense is especially significant for compatibilist RS theory, it is also a condition that is essential to any theory, including libertarian theory, that acknowledges moral sense as essential to holding agents responsible. Put more generally, no adequate theory that understands responsibility in terms of “deep assessment” and moral sense can deny that responsible agency (i.e., being responsible) also requires moral sense.

NOTES

Versions of this paper were read at the Twenty-fifth Hume Society Conference (Stirling, Scotland; 1998), the Western Canadian Philosophical Association (University of British Columbia; 1998), the Inland Northwest Philosophical Conference (University of Idaho; 1999), the University of Calgary (2002), the University of Alberta (2003), the University of Victoria (2003), and the North Carolina Philosophical Society (Duke University; 2005). I am grateful to the audiences, and especially my commentators and co-speakers on these occasions: Pat Greenspan, Joe Campbell, and Stephen Scott. Thanks also to Kari Coleman, Owen Flanagan, John M. Fischer, Walter Glannon, Robert Gunnarsson, Ish Haji, Michael McKenna, Dennis McKerlie, Jan Narveson, Geoff Sayre-McCord, and David Widerker for helpful comments and discussion.

1. P. F. Strawson, “Freedom and Resentment,” reprinted in *Free Will*, ed. G. Watson (Oxford: Oxford University Press, 1982). Hereafter abbreviated as *FR*.
2. P. F. Strawson, *Skepticism and Naturalism* (London: Methuen, 1985), esp. 38–39.
3. *FR*, 64–65.
4. *Ibid.*, 64–67.
5. On this matter, see my “Strawson’s Way of Naturalizing Responsibility,” *Ethics* 102 (1992): 287–302.
6. John Martin Fischer and Mark Ravizza, eds., *Perspectives on Moral Responsibility* (Ithaca, N.Y.: Cornell University Press, 1993), 18; their emphasis.

7. As I will explain further below, it is easy to slide from this observation to the mistaken view that moral sense has no role to play in accounting for these capacities and conditions.
8. *FR*, 75–76.
9. John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge University Press, 1998), 13–14. Hereafter abbreviated as *RC*.
10. Cp. R. J. Wallace, *Responsibility and the Moral Sentiments* (Cambridge, Mass.: Harvard University Press, 1994). Hereafter abbreviated as *RMS*. Wallace maintains that “rational self-control” requires that the agent is able “to grasp and apply moral reasons and to regulate [his] behaviour by the light of such reasons” (*RMS*, 86–87; cp. 156–61). Fischer and Ravizza take the same general view and provide a more detailed account of how this condition of rational self-control should be interpreted. These details, however, are not my concern in this context.
11. See, e.g., *RC*, 108f; 197f; 230f.
12. *Ibid.*, 208f.
13. *Ibid.*, 211.
14. *Ibid.*, 221–23; cp. Galen Strawson, *Freedom and Belief* (Oxford: Clarendon Press, 1986), esp. ch. 13.
15. *RC*, 238.
16. *Ibid.*, 236.
17. For more on this see my “Critical Notice of *Responsibility and Control: A Theory of Moral Responsibility*,” in *Canadian Journal of Philosophy* 32 (2002): 587–606.
18. *RC*, 236.
19. There are, of course, “structural” or “time-slice” responses to this line of incompatibilist criticism. For example, it may be argued that some sense can be made of how an agent can secure control over the process of acquiring reason responsive dispositions by way of second-order, critical self-reflection. The fact remains, however, that even if this were possible (and there may be skepticism about this) any further process of this kind must, ultimately, be guided by reason responsive dispositions that are unchosen and not consented to. For more of this, see my “Critical Notice of *Responsibility and Control*,” 601f.
20. Strictly speaking, on Wallace’s account all that is required for holding an agent responsible is that we believe that “such emotions would be warranted ... despite the fact that we happen not to feel them” (*RMS*, 77). However, Wallace goes on to emphasize the point that it remains true that without the relevant relationship to some reactive emotions “blame would be rendered superficial” (*RMS*, 78). Fischer and Ravizza argue, along similar lines as Wallace, that a morally responsible agent is “an apt candidate” for reactive attitudes but “need not be an actual recipient or target of any such attitudes” (*RC*, 7–8). The point that emerges out of all this is that although in particular cases we may be able to hold an agent responsible without any actual emotion or reactive attitude being involved, this cannot be the general rule or responsibility would be rendered “superficial.” Holding agents responsible is not simply a matter of having factual beliefs of some kind (i.e., as Strawson argued).
21. *RMS*, 12–15. Fischer and Ravizza are not committed to a “Kantian” view of agency.
22. Wallace explicitly allows that this can happen and cites the (fictional) example of “Mr. Spock of Star Trek fame” (*RMS*, 78 n. 41). The fact that “Spock” is a fictional character is itself significant.
23. See, e.g., Ishtiyaque Haji, “The Emotional Depravity of Psychopaths and Culpability,” *Legal Theory* 9 (2003): 65–82; and cp. Derk Pereboom, *Living without Free Will* (Cambridge: Cambridge University Press, 2001), xx–xxi. Haji endorses a Strawsonian (moral sense) view of holding responsible, whereas Pereboom rejects any view of this kind.
24. Haji, “The Emotional Depravity of Psychopaths and Culpability,” 73. In this context, Haji is responding directly to an earlier (unpublished) version of this paper, which was titled “Responsibility, Moral Sense, and Symmetry.”

25. This parallel is suggested by P. F. Strawson in his "Replies," in *Philosophical Subjects*, ed. Zak van Straaten (Oxford: Oxford University Press, 1980), 265.
26. An even more extreme form of incapacity would be if Jill had no capacity for pain in face of injury. A purely "intellectual" understanding of the fact of injury would leave her seriously impaired in her capacity to appreciate the significance of injury and be effectively motivated to respond to the threat of it.
27. One way of expressing this point is to say that Jack is not simply (systematically) token deprived of reactive attitudes but he is (constitutionally) type deprived.
28. The (useful) terminology "deep assessment" comes from Wallace. According to Wallace, blame is rendered superficial in any case where there is no recognizable connection with moral sentiments (*RMS*, 77–83; cp. 53f). See also my remarks in note 20 above.
29. See, e.g., William Damon, *The Moral Child* (New York and London: Free Press, 1988), esp. ch. 2. Damon observes that moral emotions, such as shame and guilt, "provide a natural base for the child's acquisition of moral values. As such, they both orient children toward moral events and motivate children to pay close attention to such events. These feelings provide the affective energy that motivate children's moral learning" (28).
30. See note 23 above.
31. For further details on this, see, e.g., Hervey Cleckley, *The Mask of Sanity*, 5th ed. (St. Louis, Mo.: Mosby, 1976), esp. 50–54. Cleckley makes the following observation: "Despite the extraordinary poor judgment demonstrated in behavior, in the actual living of his life, the psychopath characteristically demonstrates unimpaired (sometimes excellent) judgment in appraising theoretical situations" (346). Cleckley also points out (340–41) that the psychopath can act in accordance with moral requirements and expectations, and often does—which makes it especially difficult to detect these individuals. What does distinguish the psychopath, on Cleckley's account, is that he has "absolutely no capacity to see himself as others see him" (350).
32. See Wallace, *RMS*, 177–78; and compare John Deigh, "Empathy and Universalizability," *Ethics* 105 (1995): 743–63, who argues for a related (Kantian) view on psychopathy. Fischer and Ravizza also argue that psychopaths should be understood in terms of a failure of the reason-responsive mechanism of some kind or other (*RC*, 78–80, 82–83).
33. Antony Duff, "Psychopathy and Moral Understanding," *American Philosophical Quarterly* 3 (1977): 191–93. See also Michael S. Pritchard, *On Becoming Responsible* (Lawrence: University Press of Kansas, 1991), 50–51, who defends Duff's general position. Duff and Pritchard maintain that moral understanding requires more than "intellectual" understanding. It also requires "practical understanding," or an ability to care about some moral values, which the psychopath is unable to do.
34. See Cleckley's remarks cited in note 31 above.
35. Keeping in mind that it is not simply that this does not happen in this or that specific case; it is *constitutionally* impossible for this to happen when an agent lacks moral sense as such. In the same way, there is a difference between an agent who fails to respond to this or that moral consideration and one who is constitutionally incapable of doing so.
36. As we noted with Jill, it is the interaction between recognizing danger and feeling fear from the "inside" that makes her incapable of responding to danger in the same way as a normal person. We "over-intellectualize" the *rationality* of a normal agent in dealing with danger if we represent her as being like Jill.
37. Robert Hare suggests that "the prevalence of psychopathy in our society is about the same as that of schizophrenia, a devastating mental disorder that brings heart-wrenching distress to patient and family alike." Hare, *Without Conscience* (New York and London: Guilford, 1993), 2.
38. For a valuable discussion of this issue that is concerned with the general relevance of emotional incapacity and psychopathy, see Walter Glannon, "Psychopathy and Responsibility," *Journal of Applied Philosophy* 14 (1997): 263–75; and also Pat Greenspan, "Responsible Psychopaths," *Philosophical Psychology* 16 (2003): 417–29.

39. In this paper I have justified symmetry in terms of the particular demand of the condition of moral sense—taking moral sense to be essential for holding responsible. However, the argument for symmetry can be formulated in more general terms. The basic point is that any agent who is unable to hold himself responsible (i.e., effectively praise and blame himself and others) has a shallow and impoverished understanding of moral considerations. This incapacity affects the agent's ability both fully to understand and be motivated by moral considerations. Even a purely "intellectualist" account of what is involved in holding agents responsible should still, on this view, be committed to the general condition of symmetry.
40. Cp. Strawson, *FR*, 23. Strawson's remarks are not entirely clear on this (important) point, and this may explain why some of those who follow him seek to combine a "Kantian" conception of moral agency with a Strawsonian account of holding responsible.
41. I discuss this problem in "Selective Hard Compatibilism," in Joseph Campbell, Michael O'Rourke and Harry Silverstein, eds., *Action, Ethics and Responsibility: Topics in Contemporary Philosophy*, Vol. 7 (Cambridge, Mass.: MIT Press, forthcoming).
42. Daniel Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Oxford: Clarendon Press, 1984), 12. Although Dennett defends a version of reason responsive or rational self-control theory, he does not endorse any Strawsonian account of *holding* responsible—and therefore he is not an RS theorist.
43. The more general error here is the supposition that emotional incapacity is irrelevant to (intellectual) cognition. For an influential discussion of this general issue, see, e.g., Antonio Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (London: Papermac, 1996). In the case of moral theory, the major source of this error is found in Plato and Kant (not Descartes).