

## RESPONSIBILITY AND AWARENESS

By: [MICHAEL J. ZIMMERMAN](#)

Zimmerman, Michael J. "Responsibility and Awareness," *Philosophical Books*, 50(4) (2009): 248-261.

Made available courtesy of Wiley-Blackwell.

The definitive version is available at [www3.interscience.wiley.com](http://www3.interscience.wiley.com) or <http://dx.doi.org/10.1111/j.1468-0149.2010.00497.x>

**\*\*\*Note: Figures may be missing from this format of the document**

### **Article:**

George Sher's book<sup>1</sup> provides an elegant, concise, and subtle elaboration and defense of the epistemic component of that conception of responsibility that would appear to underlie our everyday ascriptions of responsibility. It is by far the most thorough treatment of this topic of which I am aware. In the first section of this review, I will give an exposition of the highlights of Sher's discussion. In the second section, I will undertake a critical evaluation of this discussion.

### **Part I**

The type of responsibility with which Sher is concerned is that which a person may bear for events that have occurred in the past. Such responsibility may be either moral or prudential and either positive or negative. Positive moral responsibility has to do with the praise that someone is due for having acted well in respect of doing what is morally right; negative moral responsibility has to do with the blame that someone is due for having done something morally wrong. Positive prudential responsibility has to do with the credit that someone is due for having acted successfully in the pursuit of some personally significant goal; negative prudential responsibility has to do with the criticism that someone is due for having acted foolishly. Sher focuses primarily on the negative aspect of responsibility, both moral and prudential, although in the end, he seeks a unified account of all four modes of responsibility.

Ever since Aristotle, most philosophers who have investigated the matter of what it takes for a person to be responsible for some past event have drawn a distinction between two necessary conditions of responsibility, one having to do with that person's freedom of will or action, the other having to do with his awareness or understanding of the nature of what has occurred. Sher shares this approach, calling the first condition the voluntariness condition and the second the knowledge condition. (He does so with some misgivings, recognizing that the most plausible version of the first condition might only involve a type of "origination" that falls short of voluntariness, and that the most plausible version of the second condition might only involve a type of epistemic state that falls short of knowledge.) As noted above, it is on giving a detailed account of the second condition that Sher focuses his efforts. Like most philosophers, Sher also assumes that a person may be responsible not only for the acts that he has performed, but also for the omissions that he has allowed, and, moreover, for the outcomes of his acts and omissions. For the sake of brevity, however, he tends to write exclusively of responsibility for acts, but he always has omissions and outcomes also in mind. Since he subscribes to a coarse-grained approach to event-individuation, he sometimes writes more particularly of responsibility for those features of acts (or omissions or outcomes) that serve to render them right or wrong, prudent or foolish.

In order to motivate his own account of the epistemic condition of responsibility, Sher takes as his foil a rival account that he calls the searchlight view, according to which a person's responsibility for an act (or omission or outcome) extends only to those features of the act of which he was aware—those features that were "illuminated by the searchlight of his consciousness" when he performed the act. He says that this view is both "popular" and "pervasive" (p. 9) but problematic. While he has no quarrel with the claim that one can indeed be

responsible for those features of one's act of which, at the time of its performance, one was aware, he notes that the claim that one can be responsible *only* for such features would appear to conflict with many everyday judgments, according to which responsibility is to be ascribed less restrictively. To show this, Sher presents a set of nine cases that can roughly be divided into three subsets, one having to do with involuntary lapses of judgment, the second with poor judgment, and the third with lack of moral insight. Here are representative cases from each subset (pp. 24, 26, and 28):

*Hot Dog.* Alessandra, a soccer mom, has gone to pick up her children at their elementary school. As usual, Alessandra is accompanied by the family's border collie, Bathsheba, who rides in the back of the van. Although it is very hot, the pick-up has never taken long, so Alessandra leaves Sheba in the van while she goes to gather her children. This time, however, Alessandra is greeted by a tangled tale of misbehavior, ill-considered punishment, and administrative bungling which requires several hours of indignant sorting out. During that time, Sheba languishes, forgotten, in the locked car. When Alessandra and her children finally make it to the parking lot, they find Sheba unconscious from heat prostration.

*Colicky Baby.* Scout, a young woman of twenty-three, has been left in charge of her sister's baby. The infant is experiencing digestive pains and has cried steadily for hours. Scout has made various attempts to ease its discomfort, but nothing has worked. Finally, to make the child sleep, she mixes vodka with its fruit juice. The child is rushed to the hospital with alcohol poisoning.

*Bad Weather.* It is 1968, and amerika (a nom de guerre) is a member of the Weather Underground. Sensitive and conscientious as a child, amerika has been rethinking his moral beliefs. In a series of stages, he has become convinced, first, that capitalism is deeply unjust; next, that nothing short of revolution will bring change; and, finally, that the need to rectify massive injustice far outweighs the rights or interests of mere individuals. To procure funds for the Revolution, amerika takes part in a robbery in which a bank guard is killed.

In each of these cases, Sher says, “the agent would definitely be blamed and might well be liable to punishment” (p. 24), despite being unaware of the wrongness of her or his behavior. This seems to me absolutely right as a description of the sort of reaction that the agent can expect to receive in response to what she or he has done. (Sher also notes that there are permutations on these cases in which the agent acts foolishly, rather than wrongly, and can expect to be held prudentially, rather than morally, responsible.) Since the searchlight view implies that the agent is responsible in none of these cases, it is clear that it clashes with many of our everyday judgments. So what are we to do: reject the searchlight view, or reject our everyday judgments? As Sher says, we should take the latter tack only if a case can be made for the searchlight view that is stronger than the case for retaining our current judgments. He believes that this cannot be done.

It might be thought that the conflict between the searchlight view and our everyday judgments has been overstated, in that a qualified version of the searchlight view could accommodate these judgments after all. Some philosophers have subscribed, not to the searchlight view as formulated above, but to the following position: a person's responsibility for an act extends only to those features of the act of which he is aware, *unless* his lack of awareness of these features is a consequence of some other act of whose features he was aware. The idea here is that, contrary to what the searchlight view in its original formulation says, it *is* possible after all to be negatively responsible for ignorant behavior, as long as one's ignorance can be traced to behavior of whose wrongness or foolishness one was not ignorant. But, as Sher rightly says, this qualified view affords no effective compromise in the present context, since there is no reason to believe that the agent's ignorance in the cases that have been cited can be traced to any such nonignorant origin, and yet the agent can still expect to be blamed and perhaps also punished nonetheless.

So, what case might be made for the searchlight view, qualified or unqualified? Here, Sher says, one must engage in some “imaginative reconstruction,” since advocates of the searchlight view have afforded it “strikingly little direct attention” (p. 41). He entertains two routes to the searchlight view that he finds suggested

in the writings of some of its proponents. One route, which he mines from some writings of Christine Korsgaard and Hilary Bok, concerns the idea that responsibility is a “practical concept” that plays a role only in the context of an agent’s deliberations about what to do; since an agent can only deliberate about that of which he is aware, his responsibility extends only to that of which he is aware. Sher rejects this argument on the grounds that whereas deliberation is exclusively future-oriented and first-personal, ascriptions of responsibility are neither, and so it is not true that the concept of responsibility plays a role only in the context of practical deliberation. The other route to the searchlight view that Sher discusses concerns the idea that it is unfair to hold someone responsible for what is not in his control, and a person cannot control what he does not foresee. To this he responds that although it may in a way be unreasonable to expect someone to respond to reasons of which he is unaware, it does not follow that he does not have these reasons. Moreover, agents in fact can be and often are responsive to reasons of which they are unaware. This being the case, it can be perfectly fair to hold agents responsible for failing to do what they had most reason to do, even if they were unaware that they had most reason to do it. Sher concedes that this conclusion threatens the idea that responsibility requires control, a matter to which he returns later.

Given the dearth of persuasive arguments for the searchlight view, and given its incompatibility with many of our everyday judgments of responsibility, the search is now on for an alternative view of responsibility that can better accommodate these judgments. Sher turns first to the common and plausible idea that even if the agents in the sorts of cases he cites *did not* know that they were doing anything wrong, they can properly be held responsible for their behavior because they *should have* known this. Each of Alessandra, Scout, and Amerika should have realized that she or he was wrongly putting the well-being of some innocent animal or person at serious risk, and thus each may be held responsible for having done so; the violation of some applicable standard regarding what they should have been aware of suffices, under the circumstances, for their being responsible for their behavior, even though they were unaware of the violation. Sher rejects this proposal, saying (pp. 81–2):

[A]nyone who acts wrongly or foolishly is of necessity already violating one moral or prudential standard—namely, the standard with reference to which his act is wrong or foolish. Because what we are trying to understand is precisely how such an agent can be responsible for acting wrongly or foolishly despite not being aware of violating this standard, the mere fact that he *is* violating it is evidently not sufficient to establish his responsibility. However, if the agent’s unrecognized violation of the original standard is not sufficient to tie him closely enough to his act’s wrongness or foolishness to render him responsible for it, then why should his equally unrecognized violation of any further moral or prudential or epistemic standard tie him any *more* closely to it? . . . [I]f it is problematic to hold someone responsible for violating a standard of morality or prudence whose applicability to his situation he has failed to recognize, then it will hardly help matters to add that his failure to recognize that standard’s applicability also places him in violation of a further standard, this time one which dictates responsiveness to epistemic reasons, whose applicability to his situation he has *also* failed to recognize.

So far, then, Sher has in effect argued against three theses, which can be put as follows:

- (A) Someone is negatively responsible for some act if and only if
  - (1) he satisfies any non-epistemic condition necessary for such responsibility; and
  - (2) he was aware of the act’s wrongness or foolishness.
- (B) Someone is negatively responsible for some act if and only if
  - (1) he satisfies any non-epistemic condition necessary for such responsibility; and
  - (2) either (a) he was aware of the act’s wrongness or foolishness or (b) he was unaware of the act’s wrongness or foolishness and this unawareness was a consequence of some other act of whose wrongness or foolishness he was aware.
- (C) Someone is negatively responsible for some act if and only if
  - (1) he satisfies any non-epistemic condition necessary for such responsibility; and

(2) either (a) he was aware of the act's wrongness or foolishness or (b) he should have been aware of the act's wrongness or foolishness.

In the case of Theses A and B, Sher objects to the “only if” part of the thesis but not the “if” part; in the case of Thesis C, he objects to the “if” part but not the “only if” part. What he seeks to do therefore is not to dismiss these theses wholesale, but rather to supplement them in such a way that the epistemic condition for negative responsibility is fully and properly accounted for. In his pursuit of this goal, Sher briefly considers another view of responsibility, one that has recently gained prominence and which is opposed to the searchlight view. This is the view that has come to be called attributionism. Advocated by Thomas Scanlon and Angela Smith, among others, it is the view that we can be responsible not just for acts but also for attitudes such as beliefs, desires, and emotions, even if we are unaware of the wrongness or foolishness of these acts and attitudes, as long as they reflect our rational judgments in such a way that it is appropriate, at least in principle, to ask us to defend them. It is the ability to make rational judgments that constitutes someone as a responsible person, and it is the particular contents of such judgments that are to be attributed to a particular person and which make him the responsible person he is.

Sher declares attributionism, like the other views he has considered, unacceptable, not because he believes it to be wholly misguided, but rather because he finds once again that it is unable to accommodate the full range of our everyday judgments of responsibility. Consider Alessandra. The time that she spends dealing with the problems at her children's school does indeed reflect her judgment that she has good reason to do just that, but what makes her behavior wrong—the fact that she is thereby neglecting the family pet—plays no role in forming this judgment. What Sher finds wanting in attributionism, and in all the other views he has discussed, is the key idea that a person can indeed be responsible for behavior of whose wrongness or foolishness he was unaware, as long as his lack of awareness is to be explained in terms of those attitudes, dispositions, and traits (many of them unconscious) that are *constitutive* of him and which render him capable of both recognizing and discharging the sort of moral or prudential duty that he in fact failed to discharge. The partial account of the epistemic condition of responsibility (partial, because it only deals with negative responsibility) that Sher ends up endorsing, then, is this (p. 88):

When someone performs a wrong or foolish act in a way that satisfies the voluntariness condition, and when he also satisfies any other conditions for responsibility that are independent of the epistemic condition, he is responsible for his act's wrongness or foolishness if, but only if, he either

- (1) is aware that the act is wrong or foolish when he performs it, or else
- (2) is unaware that the act is wrong or foolish despite having evidence for its wrongness or foolishness his failure to recognize which
  - (a) falls below some applicable standard, and
  - (b) is caused by the interaction of some combination of his constitutive attitudes, dispositions, and traits.

In a brief discussion at the end of the book, Sher addresses the issue of positive responsibility. The full account of the epistemic condition of responsibility that emerges is this (p. 143):

When someone performs an act in a way that satisfies the voluntariness condition, and when he also satisfies any other conditions for responsibility that are independent of the epistemic condition, he is responsible for his act's morally or prudentially relevant feature if, but only if, he either

- (1) is consciously aware that the act has that feature (i.e., is wrong or foolish or right or prudent) when he performs it, or else
- (2) is unaware that the act is wrong or foolish despite having evidence for its wrongness or foolishness his failure to recognize which
  - (a) falls below some applicable standard, and
  - (b) is caused by the interaction of some combination of his constitutive attitudes, dispositions, and traits, or else

(3) is unaware that the act is right or prudent despite having made enough cognitive contact with the evidence for its rightness or prudence to enable him to perform the act on that basis.

The final issue that Sher addresses is the relation between responsibility and control. He grants that if we understand control (as he himself is inclined to do) as requiring awareness of one's options, then on his account responsibility does not require control. But this does not mean that there is no voluntariness condition that is necessary for responsibility along with the epistemic condition that he has sought to elucidate. On the contrary, as his account of the epistemic condition itself attests, behavior for which a person is responsible must originate in the constitutive features of that person. Just how to account fully for this origination relation is a task that Sher does not undertake. He is content to note that his account of the epistemic condition is consistent with a variety of approaches, whether compatibilist or incompatibilist, that invoke such considerations as a person's character, his responsiveness to reasons, the absence of coercion or compulsion, or indeed the absence of any causally sufficient condition whatsoever.

## Part II

There is much to admire in Sher's discussion of responsibility. His defense of his own view is thoughtful and thorough, and his criticism of the views of others is in many instances subtle and astute. Moreover, there is no doubt that he is right to say that the common conception of responsibility is one according to which agents in such cases as *Hot Dog*, *Colicky Baby*, and the rest are responsible for their behavior. (By "the common conception of responsibility," I mean the understanding of responsibility that appears to underlie the everyday judgments of the general populace. I do not presume that this conception is coherent. On the contrary, I will give reasons for thinking that our everyday moral judgments are in many cases at odds with certain presuppositions about responsibility to which most of us do, and all of us should, subscribe.) Sher's account of the epistemic condition of responsibility would appear to capture this conception well, certainly better than any other account that I know of. Nonetheless, I reject his account. In this section, I will try to explain why.

First, let me note two points in Sher's favor. The first is that the searchlight view, whether unqualified as in Thesis A above or qualified as in Thesis B, certainly does not accommodate the full range of our everyday judgments of responsibility. The cases that Sher gives—*Hot Dog* and the rest—amply demonstrate this fact. The second point is that Sher is certainly right to dismiss Thesis C, the simple "knew or should have known" thesis, even though it is this thesis to which, I believe, many people appeal, whether explicitly or implicitly, when rendering their judgments in cases such as *Hot Dog*. It has always puzzled me why anyone should think Thesis C acceptable. After all, as Sher in effect says in the passage I quoted above, those who appeal to it acknowledge that one can have an excuse for wrongdoing and hence that wrongdoing does not suffice for blameworthiness. One kind of excuse that is frequently tendered is the excuse of ignorance. To accept this excuse in some circumstances but dismiss it in others, simply on the basis that, in the latter circumstances, the person in question should have known what he did not know, is bizarre; for this neglects the possibility that one has an excuse for not doing (or knowing) what one should have done (or known), the very possibility that gives rise to the question of how one can be to blame for ignorant wrongdoing in the first place! It is to Sher's credit that he clearly recognizes and exposes this problem with Thesis C.

Before I turn to my reason for rejecting Sher's account, let me raise two minor quibbles. The first concerns his characterization of the searchlight view as both popular and pervasive. This strikes me as at best misleading. I am not sure whether anyone has ever accepted the unqualified version of the view expressed in Thesis A. Certainly some philosophers, of whom I am one, have endorsed the qualified view expressed in Thesis B, but my impression is that we are decidedly in the minority among philosophers, let alone among the general populace, regarding what it takes for someone to be responsible for something, precisely because the view rules out responsibility in such cases as *Hot Dog*.

My second quibble concerns Sher's rejection of the thesis that responsibility requires control. He does so for the reason given above, namely, that control requires an awareness of one's options, whereas responsibility does not. But here I think he moves too quickly over some difficult terrain. Even if it is true, as Sher says, that "an

agent who does not realize that his act has a certain feature, or that it will issue in a certain outcome, can hardly be said to exercise control with respect to that feature or outcome” (p. 145), it nonetheless also seems true that some forms of control over an option do not require awareness of all pertinent features of that option, while others do not even require awareness of the option itself. Consider Alessandra, for example. She spent considerable time dealing with her children's problems, and wrongly so, since in so doing she unduly neglected the family pet. Was she in control of her behavior? Well, she was not aware of the wrongness of what she did, and so it may be (although in fact I have doubts about this) that, as Sher says, she did not exercise control with respect to that particular feature of her behavior. Nonetheless, she was, we may assume, perfectly aware of the fact that she was taking considerable time to deal with her children's problems, and so it seems quite reasonable to say that she was in control of her behavior in that regard. In his formal account of responsibility, Sher writes of someone's “being responsible for his act's wrongness,” but in normal parlance, we tend to talk simply in terms of someone's “being responsible for his act.” We would normally say, not that Alessandra is responsible for the *wrongness* of her taking so much time dealing with her children's problems, but simply that she is responsible for *taking* so much time to do so. Here there is no pressure to loosen the link between responsibility and control, since Alessandra was in control of that for which she is being held responsible. What of her neglect of the dog? We would normally say that she is responsible for that, too, but was it in her control? Well, certainly she was not aware of her neglect in the way in which she was aware of taking so much time dealing with her children's problems. Nonetheless, there is surely *some* respectable sense in which the neglect was in her control, since whether or not it would occur depended on what choice she would make. She did not choose to act wrongly, but she did choose, wrongly, to spend considerable time dealing with her children's problems, as a result of which she wrongly neglected the dog. She could have chosen, rightly, to spend less time on her children's problems, as a result of which she would not have unduly neglected the dog. The choice was up to her, and hence the neglect was too, even though there were aspects of her options, including but not limited to their moral status, of which she was unaware. Under the circumstances, then, it seems an exaggeration to say that Alessandra did not exercise a form of control that is relevant to, and perhaps necessary for, her being responsible for her behavior. Still, I will not press this point, since I suspect that Sher would in fact agree with what I have said, even if not with the terms in which I have said it. He is, after all, fully prepared to accept some form of “voluntariness” condition on responsibility.

Let me turn now to my reason for rejecting Sher's account of the epistemic condition of responsibility. I do so because, as indicated above, I subscribe to the qualified version of the searchlight view expressed in Thesis B. I subscribe to this thesis because I am persuaded by an argument that it is true. As noted in the last section, Sher claims that advocates of the searchlight view have paid little direct attention to defending it, and so he feels compelled to engage in some “imaginative reconstruction” on their behalf. This is disappointing, in that I myself have provided an explicit argument for the view, and Sher is apparently acquainted with this argument, since on page 8 and elsewhere he cites the article in which it first appeared. In any case, let me now give this argument in the form in which I presented it in a recent book, where I discussed a case in which one person, Alf, mistreated another person, Brenda, but was ignorant at the time that his treatment of her was wrong. Here is what I said:

Question: is Alf to blame for having [mistreated] Brenda, or does he have an excuse in virtue of his ignorance [of the fact] that his behavior was wrong? (That is: overall [as opposed to merely *prima facie*] morally wrong. This qualification will be implicit throughout . . .) Surely such ignorance does not always excuse one's behavior, since one may be culpable for one's ignorance. This is commonly acknowledged. What is not commonly acknowledged, though, is that culpability for such ignorance is rare. On the contrary, our common practice indicates that we think that such culpability is frequently incurred; for we often blame people for performing actions that were wrong (or that we take to have been wrong) on the grounds that, even if they *didn't* know that what they were doing was wrong, they *should have* known this. Many would say just this in Alf's case. I believe that this practice is misguided, however, and is itself likely to result in wrongdoing.

Here is my argument. Call the item of behavior in question *A*, and grant that (1) Alf did *A*, *A* was wrong, but Alf was ignorant of this fact since, at the time he did *A*, he did not believe that it was wrong. Now (2) one is culpable for ignorant behavior only if one is culpable for the ignorance in or from which it was performed. Hence (3) Alf is culpable for having done *A* only if he is culpable for the ignorance in or from which he did *A*. However (4) one is culpable for something only if one was in control of that thing. Hence (5) Alf is culpable for having done *A* only if he was in control of the ignorance—in particular, the failure to believe that what he was doing was wrong—in or from which he did *A*. But (6) one is never directly in control of whether one believes or does not believe something; that is, any control that one has over one's beliefs and disbeliefs is only ever indirect. Moreover (7) if one is culpable for something over which one had merely indirect control, then one's culpability for it is itself merely indirect. Furthermore (8) one is indirectly culpable for something only if that thing was a consequence of something else for which one is directly culpable. Hence (9) Alf is culpable for having done *A* only if there was something else, *B*, for which he is directly culpable and of which the ignorance—the disbelief—in or from which he did *A* was a consequence. But (10) whatever *B* was, it cannot itself have been an instance of ignorant behavior, because then the argument would apply all over again to it; *B* must, then, have been some item of behavior, some act or omission of Alf's, that Alf believed at the time to be wrong. Hence (11) Alf is culpable for having done *A* only if there was some other act or omission, *B*, for which he is directly culpable and of which his failure to believe that *A* was wrong was a consequence, and *B* was such that Alf believed it at the time to be wrong.

The picture that emerges is thus one of a chain of events or occurrences, each a consequence of its predecessors, at whose origin lies some item of behavior that Alf believed at the time to be overall wrong and for which he is directly culpable. Not at the origin, but lying somewhere further down the chain, are, first, the ignorance in or from which Alf did *A* and, second, Alf's performance of *A*. For these and other such items on the chain Alf is only indirectly culpable. We may call such a chain a chain of culpability. Now, Alf has, of course, been picked at random. What is true of him is also true of Brenda, Charles, Doris, Edward, and so on. We thus arrive at the following general thesis:

*The Origination Thesis:*

Every chain of culpability is such that at its origin lies an item of behavior for which the agent is directly culpable and which the agent believed, at the time at which the behavior occurred, to be overall morally wrong.

As far as I can tell, ignorant behavior is rarely to be traced to a non-ignorant origin. It would be surprising, for example, to find any such episode of witting wrongdoing in Alf's history to which his [mistreatment] of Brenda can be traced. In the absence of any such episode, the Origination Thesis implies that Alf is not culpable for his ignorant behavior. And there is no reason to think that Alf is atypical in this respect. Thus the Origination Thesis implies that in general culpability for ignorant behavior is rare. This may be appropriately described as a deflationary conclusion, since it implies that our common practice of frequently blaming people such as Alf for their ignorant behavior is too expansive.<sup>2</sup>

I apologize for the length of this quotation, but the argument is fairly complex, and I know of no more succinct way to present it.

Once such an argument has been given, it is, it seems to me, incumbent on anyone who wishes to defend the common view that people such as Alessandra are responsible for their behavior to identify precisely where they think the argument goes wrong. Since the argument is valid, there is only one way to do this: say which premise or premises are to be rejected, and why. In my book, I entertained reasons for rejecting each of premises (2) (4) (6), and (7) (but not [8], which I regard as analytic). I will not repeat here what I had to say there, but will simply note that (of course) I found none of the reasons persuasive. Perhaps I should have found some of them persuasive, or perhaps there are other more persuasive reasons that I failed to entertain. I will not pursue the matter here, other than to say that I cannot glean from what Sher says in his book just which premise or

premises he would reject or why. It might at first seem that he would reject premise (4), according to which culpability requires control. But I hope it is clear that in this premise, I am invoking a sense of “control” that is *not* so restrictive as to require advertence to the moral status of one’s behavior, for that would beg the very question that the argument purports to answer. Understood in a less restrictive way, the premise is one that I believe Sher would accept, since, as noted earlier, he acknowledges a voluntariness condition on responsibility.

I have characterized my argument as an argument for the qualified version of the searchlight view, but in fact this is not strictly accurate. According to what I have called the Origination Thesis (which, to be clear, concerns a kind of “origination” quite distinct from that to which Sher refers, and which I mentioned in the last section, when he discusses the voluntariness condition), culpability must be rooted in a *belief* about wrongdoing, but the thesis itself does not require that this belief be a *conscious* or *occurrent* one. Nonetheless, I do go on in the book to say that, with one possible exception, it seems that the belief in question must indeed be occurrent, rather than merely dispositional. Here is what I had to say on the matter:

With one possible exception [having to do with routine or habitual actions], if a belief is not occurrent, then one cannot act either with the intention to heed the belief or with the intention not to heed it; if one has no such intention, then one cannot act either deliberately on or deliberately despite the belief; if this is so, then the belief plays no role in the reason for which one performs one’s action; and, I am inclined to think, one incurs culpability for one’s action only if one’s belief concerning wrongdoing plays a role in the reason for which one performs the action.<sup>3</sup>

Perhaps Sher would find fault with this argument, too, but again I am not sure just where or why he would do so.

There is still another reason, however, for distinguishing between the view for which I have argued and the (qualified) searchlight view as characterized by Sher, and that is that my view concerns only culpability, which I take to be the sort of negative moral responsibility that Sher has in mind when he writes of the blame that an agent may be due for doing something wrong. My view does not extend either to positive moral responsibility or to prudential responsibility, whether positive or negative. Indeed, it may even be a mistake to say that the culpability that my view concerns is equivalent to the sort of negative moral responsibility with which Sher is concerned. Perhaps Sher’s target is broader, with culpability being only one of several kinds of negative moral responsibility of a sort for which he has sought to give an account. Let me explain.

A connection is often drawn between responsibility and what have come to be called the reactive attitudes. Many philosophers hold that responsibility consists in, or is at least correlated with, the appropriateness of being the target of such attitudes and of practices, such as punishment, that incorporate such attitudes. I accept this general thesis. What is often left unexplored, however, is the fact that there are many attitudes and practices that qualify as “reactive,” and that some of these attitudes and practices may be appropriate in some cases but not in others. This is evidence for the claim that there are different *kinds* of responsibility, even different kinds of *moral* responsibility.<sup>4</sup> Elsewhere, I have called the sort of responsibility that grounds the fittingness of reactive attitudes “appraisability,” of which I take laudability and culpability to be the positive and negative modes, respectively.<sup>5</sup> Others use different terminology. Scanlon, for example, calls the sort of responsibility that grounds the fittingness of reactive attitudes “attributability.” Now, as indicated in premise (4) of my argument above, I take freedom of will to be a necessary condition of appraisability, whereas Scanlon denies that freedom of will is a necessary condition of attributability.<sup>6</sup> It may therefore appear that Scanlon and I are at odds with one another, but I am not sure that that is in fact the case. It may simply be that we are talking about two different kinds of moral evaluability in persons rather than differing about the conditions of one particular kind. Indeed, I am inclined to think that we are talking about two different kinds.

Although the view that moral responsibility requires free will is very common, Scanlon is not the first to dispute it. Robert Adams, for example, contends that the graduate of the Hitler *Jugend* is to be blamed for his vile beliefs, no matter how he came by them;<sup>7</sup> Eugene Schlossberger likewise holds that we may be praiseworthy or



blameworthy for our commitment to certain moral values, regardless of whether they were freely chosen;<sup>8</sup> and others have made similar claims. So when Scanlon suggests, for example, that a liar is not to be excused for lying just because she is incorrigible,<sup>9</sup> he is not advancing a wholly new thesis. And, I want to say, it is a thesis we should *accept*, once its limitations have been recognized.

In order to bring out these limitations, let me consider the case of a (purportedly) “incorrigible” *non*liar. Mark Twain is reputed to have said: “I am morally superior to George Washington. He couldn't tell a lie. I can and I don't.”<sup>10</sup> I believe that this witty remark contains a very important insight. Let us suppose, no doubt falsely, that Washington was speaking the literal truth, and that he was constitutionally incapable of telling a lie. We may certainly agree with Scanlon et al. that someone who cannot commit a certain form of wrongdoing warrants some form of moral recognition, but, as Twain points out, the *kind* of recognition that is warranted differs from that warranted by someone who *can* commit the wrongdoing but deliberately chooses not to do so. So, too, we may indeed say that the youthful Nazi is in some way reprehensible for the beliefs that he holds, but, I believe, he is not to be blamed for them in the *same* way as he would be if he had formed them as a result of free reflection. In brief, the kind of moral evaluation warranted by an unfree agent differs from that warranted by a free agent, and this has implications regarding the kind of reaction warranted. So too, I would say, the kind of moral evaluation warranted by such ignorant agents as Alessandra differs from that warranted by agents who consciously do wrong. In particular, the reaction of *punishment* is, I believe, deserved only when the agent has freely and consciously done wrong. If Scanlon agrees, then we are not, or may not be, at odds after all. The kind of agent evaluability that he has in mind when he talks of “attributability” will not be exactly the same as that which I have in mind when I talk of “appraisability.”

There is a general lesson to be learned here. Many philosophers have proposed accounts of moral responsibility, and these accounts often differ in their details. Some writers hold that responsibility concerns the evaluation of an agent's character, whereas I do not. Some say that moral responsibility has to do with an agent's responsiveness to reasons, whereas I again do not. Some contend that responsibility is in part a function of authenticity, whereas once again I do not. Such examples can be multiplied indefinitely. The question arises whether these other philosophers and I disagree about the nature of moral responsibility. The general lesson is that we might be disagreeing with one another, but also we might not be; perhaps we are instead talking of different, though related, modes of moral evaluability. I suspect that whether this is so depends at least in part on whether we would agree or disagree about precisely which reactions are deserved by those whom we deem morally responsible for their behavior. If we agree about just which reactions are deserved, then that is a sign that we also agree about the kind and degree of moral evaluability at stake; if we disagree about the former, that is a sign that we also disagree about the latter. But the matter is not straightforward, since these signs are not infallible. That is, it would seem perfectly possible for two people to agree about the kind and degree of evaluability at stake but disagree about just what reactions are deserved, and also perfectly possible for two people to disagree about the former and yet agree about the latter. This being the case, the question of how to distinguish kinds and degrees of evaluability is complex.

How do these observations pertain to the (apparent) differences between Sher and myself on the nature of responsibility? Well, they might pave the way to some type of compromise between us. Sher insists that Alessandra and the other protagonists in the cases he gives are responsible for their behavior, even though they were ignorant of the fact that they were doing anything wrong or foolish, and even though this ignorance cannot be traced to some nonignorant origin. I am prepared to accept this assessment (for present purposes, at least), *as long as* the particular form of responsibility that is ascribed to these agents is not that which is correlated with their being deserving of the particular reaction of punishment. Sher says that Alessandra and the rest “would definitely be blamed and might well be liable to punishment” (p. 24). I agree with this as a description of the sort of reaction that they can in fact expect to receive. I even agree that they might *deserve* some form of blame. But I deny that they deserve *punishment*.

I am not sure to what extent Sher would go along with the compromise I have offered. I doubt that he would accept it in its entirety. It is worth noting, though, that acceptance of it even in part would tend to undermine his

overall project. The account that Sher proposes purports to provide necessary and sufficient conditions for the epistemic component not just of culpability but of negative moral responsibility generally, and, beyond that, of positive moral responsibility generally, and, beyond that, of prudential responsibility, both positive and negative. If I am right, however, that would seem too tall an order for any account of responsibility to fill (unless it were riddled with disjunctive clauses). Given its correlation with the reactive attitudes, and given the wide range of such attitudes, responsibility is simply too fragmented a concept—too multifaceted and variegated—to admit of the sort of uniform, unified treatment that Sher has sought to give it. Nonetheless, his meticulous efforts in this regard are both instructive and illuminating, and there is much to gain from a careful reading of his book, which will surely and deservedly be the focus of close attention from moral philosophers for some time to come.<sup>11</sup>

### Footnotes

1. George Sher, *Who Knew? Responsibility without Awareness* (Oxford University Press, 2009, viii + 157 pp \$99.00 hardback, \$24.95 paperback).
2. M.J. Zimmerman, *Living with Uncertainty* (Cambridge University Press, 2008), pp. 175–76.
3. *Ibid.*, p. 191.
4. The remarks that follow borrow from M.J. Zimmerman, “Responsibility, Reaction, and Value,” *Journal of Ethics*, forthcoming.
5. M.J. Zimmerman, *An Essay on Moral Responsibility* (Rowman and Littlefield, 1988), Ch. 3.
6. See T.M. Scanlon, *What We Owe to Each Other* (Harvard University Press), sec. 6.4.
7. R.M. Adams, “Involuntary Sins,” *Philosophical Review*, 94, pp. 3–31, p. 19.
8. E. Schlossberger, *Moral Responsibility and Persons* (Temple University Press), pp. 117–18.
9. Scanlon, *op. cit.*, pp. 285–86.
10. Quoted in P. van Inwagen, *An Essay on Free Will* (Clarendon Press, 1983), pp. 63–64.
11. I am grateful to George Sher for helpful comments on an earlier draft of this review.