WHO KNEW?

Responsibility without Awareness

GEORGE SHER

OXFORD UNIVERSITY PRESS

2009
ACKNOWLEDGMENTS

While I was writing this book, I discussed its topics with numerous friends, colleagues, classes, and audiences, and I am pleased to acknowledge their many contributions to the development of my ideas. Those to whom I am indebted include Nomy Arpaly, Baruch Brody, Mark Heller, Neil Levy, Michael McKenna, Donald Morrison, Nicoletta Orlandi, Bernard Rollin, Edward Sherline, Steven Sverdlik, Rachel Zuckert, and the excellent graduate and undergraduate students who were enrolled in my Spring 2008 seminar in moral psychology. I owe a debt of a different order to my wife Emily Fox Gordon, who in addition to being a constant source of encouragement is the best writer I know. Our conversations about writing, hers and mine alike, are a great source of joy to me, and I have benefitted immeasurably from them.

Portions of this book have appeared in an earlier form elsewhere. Parts of chapters 2 and 6 were published in “Out of Control,” Ethics 116 (January 2006), 285–301. Copyright © 2006 by the University of Chicago. All rights reserved. Much of chapter 4 was published as “Kantian Fairness” in Philosophical Issues 15 (2005), 179–92. © 2005 Blackwell Publishing. I thank the publishers of those journals for permission to use this material here.
In that chapter, I will defend a conception of the responsible self that stands somewhere between the maximalist view that such selves are simply full human beings none of whose physical or psychological features are any more essential than any others, and the minimalist view that when people are considered under the aspect of responsibility, they are constituted only by those features—consciousness and reason-responsiveness are the two main candidates—that are general prerequisites for it. To capture what is best about each approach, I will argue that each responsible self is constituted not only by the conscious states and rational judgments in whose absence questions about his responsibility would not arise, but also by the (presumably very substantial) subset of his other physical and psychological states whose causal interaction sustains the crucial states and judgments. Because this view of the responsible self seems independently defensible, and because it dovetails neatly with the only reconstruction of responsibility’s epistemic condition that does justice to our full range of judgments about who is responsible for what, the different elements of my proposal will turn out to hang together in a satisfying way.

A number of questions, however, remain unresolved. For one thing, because chapters 6–8 deal only with the conditions under which agents are responsible for unwitting acts that are wrong and foolish, it remains to be seen whether, and if so how, the proposed account can be extended to accommodate the very different conditions under which agents are responsible for unwitting acts that are right and prudent. A second outstanding question concerns the relation between responsibility’s epistemic and voluntariness conditions: how, if responsibility does not require awareness, can it require that the features of a person’s acts for which he is responsible be expressions of his will? Yet a third question, closely related to this one, is whether an account which implies that agents can be responsible for features of their acts of which they are unaware must abandon the familiar view that responsibility requires control. In the book’s ninth and final chapter, I shall propose answers to each of these questions. Although these answers will require some adjustment of some common beliefs, we will have to abandon only certain inessential accretions that can be traced to indefensible theoretical commitments. As compensation, we will gain a clearer view of what responsibility must be—a view that manages to acknowledge its connection to reason and subjectivity without denying that its subjects are located squarely in the natural world.

**TWO**

**RESPONSIBILITY WITHOUT AWARENESS**

My aim in this chapter is to document the striking contrast between the standard way of understanding responsibility’s epistemic condition—the interpretation I have dubbed “the searchlight view”—and our actual ground-level judgments about who is responsible for what. To make this disparity vivid, I will begin by introducing nine cases in which agents seem blameworthy, and hence morally responsible, for acts of whose wrongness they are unaware. After I discuss these cases, which will serve as my touchstone in the chapters to come, I will propose a variant of each case in which the agent acts imprudently rather than wrongly, and will generalize what has been said in a number of directions. Then, with the problem fully before us, I will argue against two attempts to avoid the difficulty by relocating the acts for which the unwitting wrongdoers and foolish agents are responsible to earlier points in their histories.

I

There are many ways in which agents can fail to recognize the morally relevant features of their potential acts. Although the categories are

1. Here and elsewhere in the chapter, I will use “acts” in a broad sense that encompasses omissions as well as positive actions.
somewhat fluid, we can usefully distinguish between cases in which an agent acts wrongly because he forgets or loses track of some crucial element of his situation, cases in which he does so because he exercises poor judgment, and cases in which the problem lies in his lack of moral insight or imagination. As a first step toward documenting my claim that the searchlight view systematically fails to capture our intuitions about responsibility, I will argue that we often hold agents responsible in all three contexts.

Here, first, are three cases in which agents seem responsible for wrong acts that they performed because they forgot or otherwise lost track of crucial elements of their situation:

1. **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 bumbling 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.

2. **On the Rocks.** Julian, a ferry pilot, is nearing the end of a forty-minute trip that he has made hundreds of times before. The only challenge in this segment of the trip is to avoid some submerged rocks that jut out irregularly from the mainland. However, just because the trip is so routine, Julian’s thoughts have wandered to the previous evening’s pleasant romantic encounter. Too late, he realizes that he no longer has time to maneuver the ferry.

3. **Caught off Guard.** Wren is on guard duty in a combat zone. There is real danger, but the night is quiet. Lulled by the sound of the wind in the leaves, Wren has twice caught herself dozing and shaken herself awake. The third time she does not catch herself. She falls into a deep slumber, leaving the compound unguarded.

In each of these cases, the agent would definitely be blamed and might well be liable to punishment. However, when Alessandra enters the school, she does not choose to forget her obligations but is distracted from them; when Julian lapses into fantasy, he does not ignore the looming rocks but fails to notice them; and when Wren falls asleep, she does not set her duty aside but ceases to be aware of it. Thus, in each case, the difficulty appears to lie not in the agent’s conscious will but in something that overtakes it.

It is of course possible to insist that being distracted, daydreaming, and falling asleep are not just things that happen to people—that if the wills of Alessandra, Julian, and Wren have been overtaken, it is only because they have allowed them to be. By reintroducing conscious volition at this slightly earlier point, we would in each case reestablish a locus of searchlight control that renders the agent’s responsibility unproblematic. However, in so doing, we would also fly in the face of common experience. We all know what it is to be assaulted by an urgent problem that drives all other thoughts from our minds; to emerge from a reverie into which we have no recollection of choosing to enter; and to have our defenses against drowsiness infiltrated by a momentary lapse in our awareness of the need to sustain them. When such things happen, there is simply no point at which we are conscious of choosing to allow them to happen. Thus, if someone were to assert that all such failures of attention are nevertheless voluntary, his claim would be implausible on its face. That claim, if not backed by some powerful independent argument—and I know of none that supports it—will not so much advance our understanding of the problem as simply define it out of existence.

**II**

Involuntary lapses of attention, though common, do not account for the majority of the wrong acts for which we hold agents responsible. Hence, if those lapses were the only counterexamples to the thesis that responsibility presupposes searchlight control, then we could preserve that thesis by simply redrawing the boundary between responsible and nonresponsible agency. This would compel us to withdraw our judgment that agents such as Alessandra, Julian, and Wren are responsible, but would leave the majority of our current attributions of responsibility intact.

But, as I have suggested, involuntary lapses of attention are not the only counterexamples to the thesis that responsibility presupposes searchlight control. No less threatening, and considerably more common, are cases in which agents act wrongly because they display poor judgment. Here again are some representative examples.
But, in fact, it does not; for by taking an agent’s responsibility to depend on whether a reasonable person would have known that his act was too risky to be permissible, we will in effect be conceding that an agent who is not in this respect reasonable may be responsible even for acts that he does not recognize as impermissibly risky. In conceding this, we will also be conceding that such an agent can be responsible for acting wrongly despite the fact that he has not consciously chosen to do so. However, that agents can be responsible for acting wrongly despite not having consciously chosen to do so was precisely the possibility that Home for the Holidays, Colicky Baby, and Jackknife were introduced to illustrate. Thus, far from defending the thesis that responsibility presupposes searchlight control against a potential set of counterexamples, the introduction of the “reasonable person” test will in effect amount to an acknowledgment that that thesis cannot be defended.  

III  

So far, we have seen that many of the acts for which we hold agents responsible can be traced either to unwilled lapses in attention or to equally unwilled failures of judgment. Even by themselves, these facts would raise difficult questions about the connection between responsibility and searchlight control. However, the questions become more serious yet when we factor in a third class of counterexamples—namely, those in which the agent willingly performs an act which is in fact

---


3. In this and the preceding paragraph, I have presupposed an objective interpretation of the “reasonable person” standard—that is, one that defines the agent’s situation without reference to his own beliefs or mental attributes. If instead the standard is interpreted subjectively, so that it takes the test for responsibility to be whether a reasonable person who shared some or all of the agent’s beliefs and dispositions would have acted as the agent did, then it may indeed preserve the principle that responsibility presupposes control. However, because the subjective version of the standard will preserve the principle by implying that Joliet, Scout, and Father Poteet are not responsible for what they have done, it will go no distance toward showing how the principle can be reconciled with the intuition that they are responsible. For discussion of the difference between the objective and subjective interpretations, see Joshua Dressler, Understanding Criminal Law, 3rd edition (New York: Lexis Publishing, 2001), 131–33 and 238–39.
wrong, but whose wrongness he does not recognize because he lacks some form of moral insight or imagination. Here again are three examples.

7. **Bad Joke.** Ryland is very self-absorbed. Though not malicious, she is oblivious to the impact that her behavior will have on others. Consequently, she is bewildered and a bit hurt when her rambling anecdote about a childless couple, a handicapped person, and a financial failure is not well received by an audience that includes a childless couple, a handicapped person, and a financial failure.

8. **Bad Policy.** Sylvain, a college professor, is empathetic to a fault. He identifies readily with troubled students and freely grants their requests for opportunities to earn extra credit. Because he enters so completely into each interlocutor’s perspective, he often forgets that there are other, less aggressive students who would eagerly welcome the same chance. As a result, his grading policy is inconsistent and unfair.

9. **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.

Of these three agents, Ryland is too self-centered to recognize her anecdote as hurtful, Sylvain is too focused on the individual before him to realize that he is being unfair to others, and Amerika has taken a disastrously wrong turn in working out the implications of his moral beliefs. Nevertheless, although none of the agents willingly acts wrongly, all three seem blameworthy and one (Amerika) also deserves to be punished.

We have now encountered a total of nine cases in which agents seem responsible for wrong acts whose wrongness they did not recognize. Although the agent’s lack of awareness is crucial to all nine cases, it does not always take the same form. In three cases (**Hot Dog, On the Rocks, Caught off Guard**), the agent does not even realize that he is in a situation that calls for action, while in the remaining six he does realize this but lacks an accurate appreciation of what he ought to do. In one of these six cases (**Home for the Holidays**), the agent’s cognitive defect is due to a distorting emotion (panic), in two others (**Bad Joke Bad Policy**), it can be traced to his insensitivity to a morally relevant factor, and in the remaining three (**Colicky Baby, Jackknife, Bad Weather**) it is a product of unadorned poor judgment. In two of the latter cases (**Colicky Baby, Jackknife**), the defective judging occurs when the agent is assessing the facts, while in the third (**Bad Weather**), it occurs when he is thinking through his moral beliefs. Because these patterns of error are quite diverse, and because each one is instantly recognizable, the range of counterexamples to the searchlight view can already be seen to be broad.

IV

But acts that are morally wrong (or, on the positive side, morally admirable or supererogatory) are not the only ones for which we view agents as responsible. We also frequently view agents as responsible both for their foolish or imprudent choices and for acts that are especially clever or effective. Because the searchlight view is generally held to apply in these contexts no less than in contexts of immorality, the next thing to note is that each of our nine cases can easily be transformed from a counterexample to the claim that: moral responsibility presupposes searchlight control into a counterexample to the corresponding claim about nonmoral responsibility.

To transpose each example to the key of nonmoral responsibility, we need only replace each morally wrong act with a suitably imprudent one. To effect the requisite changes, let us now suppose that the hot dog of which Alessandra loses track during the wrangle at school is not Sheba, but only a part of that evening’s dinner that begins to burn and causes a kitchen fire; that the boat that Julian runs onto the rocks is not a ferry but his treasured ChrisCraft; and that what Wren sleeps through is not a period of danger to her comrades but an appointment upon which her own promotion depends. Let us suppose, further, that the person whom the panicky Joliet shoots is not her son but herself; that what Scout fails to consider is not how alcohol will affect her young niece but how she herself will be affected by ingesting the contents of her sister’s pharmacopia; and that instead of cutting in front of the eighteen-wheeler, Father Poteet swerves sharply to the right and plunges into a drainage ditch. Let us suppose, finally, that Ryland is summarily fired when word of her tasteless anecdote reaches her
employer; that the upshot of Sylvain’s susceptibility to special pleading is a flood of other special pleaders; and that as amerika ages, his misguided beliefs about the moral insignificance of the individual are replaced by equally misguided beliefs about the spiritual benefits of self-starvation and daily colonic irrigation.

Although these changes do not eliminate the wrongness of all nine original acts—at least where Ryland and Sylvain are concerned, the original act remains the same—they do transform each case into an instance of self-inflicted harm. Moreover, in each amended case, we may assume both that the agent did not realize that what he was doing would harm him and that his failure to realize this had the same source as his failure to recognize his act’s wrongness in the original example. Under these assumptions, the searchlight view will imply that none of the nine agents in the amended cases are responsible for the harms or disadvantages that have resulted from their imprudence. However, when we think carefully about these agents, we encounter a number of reactions that strongly suggest that we do consider them responsible.

One such reaction concerns the urgency of mitigating the damage the agents have done to themselves. Treating Joliet’s bullet wound, Scout’s drug overdose, Father Poteet’s internal injuries, and amerika’s damaged digestive system may be very costly and may require the use of resources (organs for transplant, sophisticated diagnostic machinery) the demand for which far outruns the supply. Even if it is agreed that all citizens should be provided with medical care, many will feel that Scout and amerika, and perhaps also Joliet and Father Poteet, should be relegated to a lower place in the queue of claimants than those who are not at all responsible for their condition. Given (what I take to be) the pervasiveness of this reaction, we evidently view these agents as responsible for their imprudently self-inflicted harms despite their failure to foresee those harms.

And, along similar lines, I suspect that many would be less willing to make personal sacrifices, and would be willing to make fewer such sacrifices, to help agents like Julian, Wren, and Sylvain than to help others with identical needs but different histories. Are we as willing to wade into rocky water to retrieve a boat that was breached because its owner was not paying attention as we are to retrieve a boat that was damaged by a violent sudden storm? Would we be as ready to make a special trip to meet with someone who slept through an appointment as we would to meet with someone whose car broke down? Are we as willing to do part of the work of a colleague who is swamped because he didn’t think things through as we are to do the work of someone who was incapacitated by illness? If, as I suspect, the answer in each case is “no,” then the reason is again that we regard the first member of each pair, but not the second, as responsible for his own predicament.

Although these reactions suggest that we regard the majority of the agents in our amended cases as responsible, we may not have the reactions about all nine cases. To cite what look like the two most refractory examples, I would expect many to feel that Father Poteet has as strong a claim to be provided with scarce medical resources as anyone else and that helping Alessandra is no different from helping someone whose kitchen caught fire because the wiring was faulty. Yet even if we do view Father Poteet and Alessandra as having strong claims to be helped, I think most of us would also feel that Father Poteet has good reason to reproach himself for his poor judgment and that Alessandra has good reason to reproach herself for forgetting what was on the stove. In this crucial respect, Father Poteet remains very different from a driver whose injuries are due to mechanical failure or a fainting spell and Alessandra remains very different from someone whose kitchen caught fire because the wiring was faulty. Thus, even in the most problematic of our cases, our reactions suggest that we regard the agents as responsible for the self-inflicted harms that they failed to foresee.

V

Even when they are expanded to encompass acts that are imprudent as well as wrong, my nine examples remain a very limited sample of the types of cases in which we take agents to be responsible. For this reason, there may still be questions about how commonly we are willing to hold agents responsible for acts of whose wrongness or foolishness they are not aware. To bring out the pervasiveness of this phenomenon, I want next to call attention to some of the ways in which the examples can be generalized.

It should be clear, first, that each type of failure that I have mentioned—forgetting, bad judgment, insensitivity—can have many different causes and can take many different forms. To cite just a few
possibilities, a person may fail to realize that he has a certain obligation (to return a phone call, look into the causes of his child’s bad grades, break a piece of bad news tactfully) because his head is abuzz with information, because he finds it unpleasant to think about the situation that is generating the obligation, because he is preoccupied with his team’s playoff chances, or for any number of other reasons. A person may fail to plan prudently, or may fail to see how to execute his well-wrought plan, because he is in a hurry or overeager to please, because he is worried about letting a golden opportunity slip through his fingers, or simply because he is not good at this sort of thing. A person’s moral or prudential judgment may be clouded by anxiety, exhilaration, depression, anger, grief, or any other strong emotion. A person’s thinking may be so rigid and stereotyped that the natural thing to do, or a tactful or rhetorically effective way of saying what must be said, just does not occur to him. And each such pattern can itself interact with any of the others, as when an agent’s unwilled lapse of attention is itself what leads him to judge poorly or to fail to recognize a morally relevant consideration. For each of the possibilities just listed, we can easily tell many stories that would lead most people to judge that the agent is responsible for the ensuing wrong or foolish act.

So far, I have concentrated on cases in which agents seem responsible for acts whose morally or prudentially relevant features they do not recognize. However, an even more common type of case is that in which the agent is aware of all the morally and prudentially relevant features of his situation, but is mistaken about their moral or prudential weight. A wrongdoer who underestimates the weight of a given moral consideration stands somewhere between a psychopath, who doesn’t recognize that consideration as relevant at all, and an akatic, who is fully aware of its true weight but is not (or not sufficiently) moved by it. Agents of this sort are counterexamples to the searchlight view because they are responsible despite the fact that their inaccurate weighting of the reasons renders them unaware of the wrongness or foolishness of what they are doing.

Here as above, examples are easy to come by. We are all familiar with the agent who realizes that he is doing something unsavory but incorrectly takes his act to be justified by its good consequences; with the student who realizes that cheating on exams is wrong but thinks it cannot be too wrong if so many others are also doing it; and with the meddler who sees value in respecting another’s autonomy but who too readily believes that it is more important to prevent some harm. Along similar lines, we all know people who realize that certain acts are risky, but who fail to realize that they are acting foolishly because they underestimate the importance of what they are risking, because they cannot vividly imagine certain possibilities, or because they just don’t care what will happen when they are old. In these sorts of cases, even more clearly than in the others, an agent’s failure to recognize the wrongness or imprudence of what he is doing is generally thought not to relieve him of responsibility for doing it.

Given the variety of contexts in which agents can be responsible despite their failure to realize that they are acting wrongly or foolishly, and given easy availability of plausible illustrations, I suspect that clear-eyed decisions to act wrongly or foolishly are actually pretty rare. However, for present purposes, I need not say anything this strong. Instead, to propel my argument forward, I need only point out that whatever the exact ratio of unwitting to witting instances of wrongdoing and foolishness within the domain of apparently responsible action, that proportion is far higher than the searchlight view can allow.

VI

We have seen that agents often seem responsible for acts of whose wrongness or foolishness they were not antecedently aware. But does this really compel us to choose between rejecting the searchlight view and drastically revising our attributions of responsibility? Can’t a proponent of the searchlight view somehow square the fact that none of our nine agents was aware that he was acting wrongly or self-destructively with the intuition that each agent is responsible for what he did?

4. Although Michael J. Zimmerman is dealing with only a relatively narrow range of cases, he draws a conclusion of just this sort when he writes that “the conditions [for culpable ignorance] are pretty restrictive,” and that therefore “culpable ignorance occurs less frequently, perhaps far less frequently, than is commonly supposed.” (Michael J. Zimmerman, “Moral Responsibility and Ignorance,” Ethics 107 [April 1997], 411). For further discussion along these lines, see Gideon Rosen, “Culpability and Ignorance,” Proceedings of the Aristotelian Society 103 (2003), 61-83.
According to many philosophers, we can indeed achieve a reconcili-
cation by relocating each agent’s conscious choice to some earlier
point in his history. If an agent previously realized that what he was then
doing would subsequently lead to his acting wrongly without realizing it
(or, less exactly, if he realized that what he was then doing would
subsequently raise the likelihood of his acting wrongly without realizing it),
then his responsibility for his current wrong act may simply be a
consequence of his responsibility for his earlier one. Hence, to complete
my demonstration that the searchlight view is massively incom-
patible with our beliefs about who is responsible for what, I must
explain why this maneuver fails. To keep things brief, I will confine
my discussion to the original versions of the nine examples. However,
everything I say will also apply, mutatis mutandis, to their nonmoral
variants.

The suggestion that what makes an agent responsible for an unwitting
wrongful choice is always some previous choice that was witting
as well as wrongful is already familiar. We encountered a version of
it when we asked whether Alessandra had chosen to allow herself to
be distracted, whether Julian had chosen to-daydream, and whether
Wren had chosen to let herself fall asleep. In this initial version, the
suggestion was found not to be credible. However, even if we cannot
locate the fateful choice at the moment immediately before the agent
becomes incapable of recognizing his act as wrong, we may still be able
to locate it at some yet earlier moment.

Interestingly, proposals of this sort can be extracted from two quite
different bodies of literature. One, predictably, is the literature on
culpable ignorance. It is often said that a necessary (though not a
sufficient) condition for culpability for an act that one does not recog-
nize as wrong is the prior wrongful performance of what Holly Smith
has called a “benifiting act”—that is, an act in which the agent “fails to
improve (or positively impairs) his cognitive position” in a way that
creates a risk of precisely the type of wrongdoing that later occurs.
Taking our cue from this, we might suppose that the wrongful choices
that render the agents in our examples responsible consist precisely of
such benifiting acts. The other body of literature that might be adapted
to provide a way of relocating the moment of knowing wrongful choice
is the literature on character formation. Inspired by Aristotle, many
philosophers have pointed out that we are at least to some extent able
to shape our characters—that, among other things, we become good by
practicing virtue and avoiding vice, bad by doing the reverse. This
suggests that the wrongful choices that have rendered our agents
responsible may be those through which they solidified the habits that
subsequently prevented them from realizing that what they were doing
was wrong.

Because there are always innumerable past acts and omissions but
for which a wrongdoer would not have failed to realize that he was not
acting as he should, each of these proposals has some initial credibility.
However, problems emerge as soon as we try to flesh the proposals out.
One obvious problem is that in the majority of our nine examples, the
agent’s lack of awareness that he is acting wrongly simply does not
appear to be traceable to any prior wrongful act or omission. A further
problem is that even where there has been previous wrongdoing, the
agent’s responsibility does not appear to depend on his previously
having been aware that he was acting wrongly. For ease of exposition,
I will develop these objections first as they pertain to wrongful benifiting
acts and then as they pertain to wrongful failures to develop good
habits and traits.

Have the agents in our examples knowingly and willingly per-
formed wrongful benifiting acts? To see that the majority have not,
we need only remind ourselves of how unpredictable their situations
were. For example, when Alessandra arrived at the school, the dispute
that she encountered was not one that she could have anticipated.
Because she had no previous reason to expect to be distracted, she

5. Holly Smith, “Culpable Ignorance,” The Philosophical Review, 92 (October 1983),
547; see also Zimmerman, “Moral Responsibility and Ignorance.”

6. For relevant discussion, see Jonathan Jacobs, Choosing Character: Responsibility for
Virtue and Vice (Ithaca: Cornell University Press, 2001); Gregory Triantofy, “Natural
Affection and Responsibility for Character: A Critique of Kantian Views of the Virtues,” in
Owen Flanagan and Amelie Oksenberg Rorty, eds., Identity, Character, and Morality
(Cambridge, MA: Bradford, 1990), 93–109, and Gary Watson, “Skepticism About Weak-
ness of Will,” in his Agency and Answerability: Selected Essays (Oxford: Oxford University
also had no previous reason to take precautions against being distracted. Thus, on the most natural reconstruction of Hot Dog, there was no previous point at which Alessandra negligently failed to do something that would have prevented her from forgetting Sheba. Nor, similarly, does the most natural reconstruction of Jackknife involve any wrongful benigning act on Father Poteet's part; for because he had no reason to mistrust his judgment, he also had no previous reason to take precautions against its failing. And, although I shall not bother to argue the point, I think we would be just as hard-pressed to attribute wrongful benigning acts to Julian in On the Rocks, to Wren in Caught Off Guard, or to Amerika in Bad Weather.7

The other four examples are different. It is clear enough that Joliet should have investigated before shooting the intruder and that Scout should have tried to find out whether babies can tolerate alcohol. It may also be true that Ryland should have stopped to think before telling her anecdote and that Sylvain should have reflected before agreeing to allow his student to earn extra credit. Hence, when it comes to these cases, the attempt to relocate the moral lapse does have some initial credibility.

However, even if all four unwitting wrong acts can be traced to previous wrongful benigning acts, it does not follow, nor is it plausible to say, that any of these agents must have recognized the previous act as either benigning or wrongful. Given Joliet's panic, it seems unlikely that she even thought of investigating the identity of the intruder, much less realized that she was acting wrongly by not doing so. Nevertheless, even if she never was aware that she was not acting as she should, she remains responsible for shooting her son. Mutatis mutandis, the same is true of Scout in Cokiey Baby, of Ryland in Bad Joke, and of Sylvain in Bad Policy. Thus, even if we assume that each agent did previously perform a wrongful benigning act, the four cases will remain counter-examples to the claim that agents are not responsible for their unwitting wrong acts unless those acts can be traced to previous benigning acts which in their turn were unfitting as well as wrongful.

So far, I have argued only that we cannot accommodate our intuitions about the nine cases by saying that what each agent is really responsible for is an earlier wrongful benigning act. But neither, similarly, can we accommodate those intuitions by saying that what each agent is really responsible for is an earlier wrongful failure to cultivate habits or traits that would subsequently have prevented him from acting wrongly. This suggestion, too, fails partly because most of our cases are best interpreted as involving no such failures and partly because even the remaining cases do not compel us to assume that the relevant failures were unfitting as well as wrongful.

To see that the majority of our cases involve no previous wrongful failures to cultivate good habits and traits, we must note, first, that agents are at best obligated to prevent the development of habits and traits that are markedly worse than normal. We may indeed act wrongly when we allow ourselves to develop vices, but we do not act wrongly by not trying to become moral saints. This is significant because in the majority of our examples, the relevant wrongful acts are not best understood as stemming from habits or traits that are markedly worse than normal. There is no particular reason to suppose that Alessandra, Wren, and Julian are any more irresponsible than the average person or that Joliet and Father Poteet are unusually indifferent to the safety of others. Because we need not suppose that any of these agents has been remiss in not developing his character, we draw a blank when we look for wrongfully foregone opportunities for self-improvement in which to locate their responsibility for their later wrong acts. Should Alessandra have been taking Gingko Biloba to improve her memory? Was Father Poteet remiss in not sharpening his reflexes on video games? Should Joliet have been on Paxil?

7. For a further illustration of the problems associated with this approach, consider a case that is introduced by Eugene Schlossberger in his book Moral Responsibility and Persons (Philadelphia: Temple University Press, 1992). Schlossberger asks us to consider Joel, an American in England who causes a traffic accident because he does not realize that the British drive on the left. To explain why Joel is responsible for causing the accident, Schlossberger writes that "other things being equal, Joel is responsible for instantiating the trait of not having bothered to check the local traffic laws before driving... For had he placed more importance on [the safety of other motorists], he would have thought that the risk to their safety, though, as far as he knew, rather slight, outweighed the extra three minutes of sleep he gained by not getting up earlier to look up the law, to make sure that no curious but important feature of English traffic laws escaped his attention" (109). This passage is both typical of what those who wish to associate all instances of culpable ignorance with prior benigning acts must do and so far-fetched that it nicely illustrates the reasons for not adopting their strategy.
Here again, the remaining cases are different. As described, Scout sounds unusually irresponsible, Ryland sounds unusually insensitive, Sylvain sounds like he lacks a basic sense of fairness, and amerika sounds like he has lost his moral compass. Hence, in Colicky Baby, Bad Joke, Bad Policy, and Bad Weather, there is at least theoretical room for the claim that what the agent is really responsible for is wrongfully having allowed himself to develop the vice that his unwitting wrong act manifests.

Yet even if we waive the objection that each incipient vice was itself a barrier to the agent’s recognition of the need to take steps to prevent its further development—the objection that, for example, the very self-involvement that now makes Ryland so insensitive is also likely to have made her unaware of the need for sensitivity training—this strategy for relocating what the agents are really responsible for will remain far-fetched. The basic problem is the relation between the way a person lives and the character he will end up having is generally transparent only in retrospect. Our characters develop slowly and by accretion, and their development is influenced not only by the decisions we make and the situations into which we enter, but also by our innate tendencies and the innumerable unchosen exigencies with which life presents us. Given the complexity of each factor, and given the exponentially greater complexity of the ways in which the different factors can interact, we rarely make decisions with the clear understanding that they will cause us to acquire traits or habits that are markedly worse than normal. Although there may be a handful of activities that can be predicted to have such effects—working in a missile silo and being a prison guard are two that come to mind—our view of the future is usually too clouded to license such inferences. We generally have little idea of which traits we will develop if we do or do not marry a certain person, undertake a certain career, or put down roots in a certain part of the country. A fortiori, we are generally not in a position to appreciate the cumulative effects of the innumerable smaller decisions of which our daily lives consist. Thus, even if we grant that Ryland, Scout, Sylvain, and amerika were remiss in not taking what in retrospect can be recognized as the steps that were necessary to prevent the development of their current bad traits, it will remain implausible to suppose that their wrongful failures to take those steps were themselves witting.  

8. For additional defense of the view that we rarely have an unobstructed view of the ways in which our present actions will affect our future characters, see Nomy Arpaly, Unprincipled Virtue: An Inquiry into Moral Agency (Oxford: Oxford University Press, 2003), 139–44.


10. For further criticism of the “tracing” approach to responsibility in cases of ignorance or lack of control, see Manuel Vargas, “The Trouble With Tracing,” Midwest Studies in Philosophy 29 (2005), 269–91; Angela Smith, “Responsibility for Attitudes: Activity and Passivity in Mental Life,” Ethics 115 (January 2005), 236–71; and Michael McKenna, “Putting the Lie on the Control Condition for Moral Responsibility,” Philosophical Studies 139, 29–37. In their recent Essay “The Truth About Tracing” (Nous, forthcoming), John Martin Fischer and Neal Tognazzini defend the tracing approach against a variety of counterexamples proposed by these authors.