

BLAMING FOR UNREASONABLENESS ACCOUNTABILITY WITHOUT ILL WILL

Alisabeth Ayars

MANY THEORISTS of moral responsibility endorse the principle that X is blameworthy for a wrong act A only if in doing A , X expressed ill will. Call this the *Quality of Will Condition*. Ill will is, roughly, an objectionable lack of concern for morality or the morally significant interests of others. Exactly what ill will consists in is controversial. All hands agree that ill will need not involve malice, so the term is something of a misnomer: it would be better to speak of an “insufficiently good will” or perhaps an “objectionable pattern of concern.” However, nothing I say in this paper turns on any particular view of ill will, other than that ill will is a matter of what the agent wants or cares about.

On one prominent conception, an agent’s quality of will is a matter of her responsiveness to moral reasons.¹ A goodwilled person has a final desire to take courses of action that have right-making features (like kindness or fairness), and a final desire to avoid actions with wrong-making features (like cruelty). In Arpaly’s words: “To say that a person acts out of moral concern is to say that a person acts out of an intrinsic (noninstrumental) desire to follow (that which in fact is) morality.”² And a person acts from ill will if she is either insufficiently responsive to moral reasons or responsive to sinister reasons—“reasons which conflict with morality.”³ A person who tortures a puppy for fun expresses ill will because she desires the puppy’s pain, and is thus responsive to a consideration that makes the act wrong. A person who fails to aid someone who is suffering might express ill will not because she is responsive to a sinister reason but because she is insufficiently responsive to the moral consideration at hand—the person’s suffering.

Proponents of the *Quality of Will Condition* sometimes argue that it is implicit in our ordinary practice of blaming. Strawson famously claimed that the

1 Arpaly, *Unprincipled Virtue*.

2 Arpaly, *Unprincipled Virtue*, 84.

3 Arpaly, *Unprincipled Virtue*, 79.

reactive attitudes—resentment, indignation, gratitude, and the like—are always responses to the good or ill will displayed by others:

The personal reactive attitudes rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves; or at least on the expectation of, and demand for, an absence of the manifestation of active ill will or indifferent disregard.⁴

Indeed, the Quality of Will Condition appears to explain central features of our practice, most notably the range of excuses we accept as grounds for withholding blame for wrongful action. It is a central feature of our blaming practice that we withhold blame when we take the agent to have acted under duress or from certain forms of blameless ignorance. Any theory of blameworthiness must explain why these excuses work when they do, and for the Strawsonian the explanation is straightforward: these excuses negate blameworthiness (when they do) by indicating that the violation was not sourced in any improper concern for the rights or interests of others. It is important to note that the Quality of Will Condition only specifies that ill will is a *necessary* condition of blameworthiness. It is consistent with thinking there are other requirements, e.g., that the agent acted wrongly, or that she possessed the capacity to recognize and respond to moral reasons.⁵

I will first argue, against this near consensus, that the Quality of Will Condition is in fact false to ordinary moral practice.⁶ This is because ordinary practice licenses blame for agents who act wrongly from epistemically unreasonable ignorance of the wrong-making features of their act, even when the act does not express ill will. This should be especially concerning to Strawsonians, who think the norms implicit in our practice are the inevitable starting point for discerning the real norms of blameworthiness.

Second, I provide a theory of culpable ignorance (and blameworthiness more generally) on which ill will is not necessary for blameworthiness. The theory is an answer to the question: In what might the culpability of unreasonable agents be grounded, if it is not ill will? There is already an answer to this question

4 Strawson, *Freedom and Resentment and Other Essays*, 347.

5 Proponents of the Quality of Will Condition include Strawson, *Freedom and Resentment and Other Essays*; Arpaly, *Unprincipled Virtue*; Rosen, “The Alethic Conception of Moral Responsibility”; Wallace, *Responsibility and the Moral Sentiments*; McKenna, “Directed Blame and Conversation”; and Graham, “A Sketch of a Theory of Moral Blameworthiness.”

6 Baron is a fellow traveler in this position in “Culpability, Excuse, and the ‘Ill Will’ Condition.”

in the literature: capacitarianism, which says that an agent is culpable for fully unwitting wrongdoing insofar as she had, but failed to exercise, the *capacity* to be aware of the wrongness or wrong-making features of her act. But all capacitarians need to distinguish between capacities that are relevant to responsibility and capacities that are not, as will become clear. The present approach draws this line in a principled way.

My theory, which I will call *Rational Capacitarianism*, entails that when an agent acts wrongly from factual ignorance, she is blameworthy only if her ignorance is caused by a failure to exercise a *rational capacity*—a capacity to recognize reasons for believing and desiring and to be generally responsive in one's beliefs and desires to one's assessment of reasons. I argue that failures of rational capacities are relevant because only such failures can be attributed to the agent in a distinctive way. My theory shows that we do not need to choose between a quality of will view (which seems too lenient) or a standard capacitarian view (which seems too harsh).

A few preliminaries: I will consider an agent to be blameworthy for an act if she is liable to certain "hot" emotional responses for it, like resentment and indignation. This follows Strawson in reducing questions of blameworthiness to questions about the appropriateness of the negative reactive attitudes.

I will deviate from the literature in focusing on ignorance that results from epistemic unreasonableness, rather than lapses of memory or inattention. The literature on negligence tends to spotlight forgetful or inattentive agents, e.g., a parent who forgets about a child in a hot car, or a driver who does not check the rearview mirror and backs into a pedestrian. But I think it is a mistake for ill-will skeptics to emphasize these cases, because it is actually *less* apparent that forgetful or inattentive agents are culpable than it is that *unreasonable* agents are culpable. This will become clearer when I explicate the notion of a rational capacity.

1. EPISTEMICALLY UNREASONABLE IGNORANCE

There are two relevant kinds of ignorance: factual and moral. A factually ignorant agent is unaware that her act has certain empirical features that make it wrong (like being risky or harmful); a morally ignorant agent may be aware of these features, but ignorant that her act is wrong (or cruel or unfair). I will focus on factual ignorance rather than moral ignorance, because it is more difficult to see how an agent whose pertinent ignorance derives entirely from factual ignorance might be culpable.⁷ Factual ignorance is often a complete excuse. Suppose *X*

7 There is an obvious way in which morally ignorant agents can still express ill will, by being insufficiently responsive to the reasons that *make* their act wrong. A Nazi who tortures

poisons *Y* by putting a substance in her coffee. If *X* innocently believed that the substance was sugar (perhaps because it was mislabeled, and she had no reason to doubt the labeling), then *X* is excused—she is not blameworthy for the wrong she committed.⁸ Strawsonians explain this in a straightforward way: if *X* was ignorant of the harm she would cause and acted only because she was ignorant in this way, the wrong was not sourced in any improper concern.

Factual ignorance is only an excuse, of course, when it is not the foreseeable upshot of some prior blameworthy act that expressed improper concern.⁹ If to relieve the boredom of a long drive I ingest a mind-altering drug and hit a pedestrian whom I do not notice crossing the street, it is no excuse that I did not know he was there. Since I am blameworthy for recklessly taking the drug, I am blameworthy for my ignorance.

Ignorance may also not excuse if it is *motivated* by ill will, as when a racist's racial animus leads him to falsely believe a Black candidate is unqualified for a job. If an agent acts from motivated ignorance, she acts from improper concern, since her ignorance expresses improper concern. Call ignorance that satisfies neither of these conditions—that is neither motivated by ill will nor the upshot of prior blameworthy conduct—*morally blameless* ignorance (leaving open that it may be criticizable in some other way, e.g., epistemically criticizable). In what follows I set aside morally blameworthy ignorance. The plan is to focus on *blameless ignorance* and to ask when and why it amounts to an excuse.

Most theorists hold that acts performed from blameless factual ignorance are always blameless. For instance, Rosen says:

If [the agent] has been neither negligent nor reckless in the management of his opinion, then his ignorance is blameless and so is the act done from ignorance.¹⁰

innocent prisoners and is fully convinced he is doing the right thing is still insensitive to the suffering of the prisoners. This is why many quality of will theorists do not think *moral* ignorance is exculpating.

- 8 In “Does Moral Ignorance Exculpate?” Harman points out that it is not ignorance that is potentially exculpating, but false belief. If you have a 0.5 credence that a substance is poison and a 0.5 credence it is sugar, then you are ignorant that it is poison, but this does not get you off the hook for poisoning someone. You are only off the hook if you have the false belief that the substance is *not* poison. This point is well taken. “Ignorance” of *p* in this paper should be construed as shorthand for false belief that not-*p*; further qualifications may also be needed. For discussion, see Harman, “Does Moral Ignorance Exculpate?”
- 9 Rosen, “Culpability.”
- 10 Rosen, “Culpability,” 63. By “neither negligent nor reckless,” Rosen means that the person's ignorance is not the upshot of a previous negligent or reckless *act*. It does not cover pure

And Zimmerman says, “culpability for ignorant behavior must be rooted in culpability that involves no ignorance.”¹¹ Indeed, it may seem almost a truism that an agent is not blameworthy for a wrong act if she was blamelessly unaware of its wrong-making features. Consider this case of Arpaly’s: Boko Fittleworth (a character in a P. G. Wodehouse novel) captures a man he finds hiding in his future father-in-law’s garden shed, thinking he is a burglar.¹² In fact, he is not a burglar; his presence is part of a secret plot in which the future in-law is willingly participating. Boko is, obviously, completely blameless for this honest mistake.

But Rosen’s view is not as obvious as it may first appear. We must distinguish *epistemically reasonable* from *epistemically unreasonable* ignorance. The cases that support the view that blameless ignorance is always exculpating tend to be cases of epistemically justified ignorance: cases in which the agent drew sound conclusions based on the evidence available to her. I am clearly off the hook for poisoning you if my belief that the substance was sugar was not only unsourced in ill will, but completely reasonable (for instance, because the canister was labeled “sugar” and I had no reason to think this was misleading). But what if the false belief from which an agent acts is, though blameless, highly irrational? Suppose the canister was labeled “cyanide—not for consumption!” but I unreasonably, though innocently, assumed it was a joke. Am I still off the hook?

Now, irrationality is sometimes the result of underlying bad will. I wish to set these cases aside (since they are instances of *motivated* ignorance), but since it may be hard to distinguish them from cases of innocent irrationality, it is worth saying a bit more about them. In *Unprincipled Virtue*, Arpaly asks us to consider someone who claims “I hate him because he is disgusting.” We are likely to infer that the speaker “puts the cart before the horse”: it is the hate that motivates the belief that the person is disgusting, not vice versa.¹³ Consider a young academic who irrationally believes that women are awful at abstract thought, despite being surrounded by brilliant female colleagues. Again, we would reasonably surmise that the man’s irrational belief is rooted in some unsavory motive, perhaps a discomfort with women’s presence in academia. These cases of *motivated* irrationality or *motivated* ignorance are clearly possible. And when an agent acts from motivated ignorance, she acts from ill will in a straightforward sense, because her ignorance itself is the result of a bad desire.

The cases that interest me are not like this. They are ones in which the agent’s

epistemic irrationality, which will be my focus. Rosen’s view is actually more extreme, as he thinks that blameless moral ignorance is also always excusing. But we can set this aside.

11 Zimmerman, “Moral Responsibility and Ignorance,” 417.

12 Arpaly, *Unprincipled Virtue*.

13 Arpaly, *Unprincipled Virtue*, 105.

irrationality is unmotivated. It is clear that human beings are sometimes just poor thinkers. As Kahneman and Tversky have shown, we are subject to a range of cognitive biases.¹⁴ For instance, we tend to be bad at reasoning about probabilities because we neglect base rates and rely too heavily on examples that immediately come to mind (the “availability heuristic”). And sometimes, of course, people simply form beliefs that are incongruent with the overall balance of their evidence, even when they are not affected by a particular bias.

Given the prevalence of garden-variety irrationality, it is clearly *possible* that someone might act badly from ignorance that reflects unmotivated irrationality but that is not further sourced in some improper pattern of concern. Whether or not such cases are common is, of course, another matter. What I hope the reader will appreciate at this point is that such cases are *possible*, and that the right theory of blameworthiness should properly classify them. Are agents who act wrongly from *unmotivated* irrational ignorance culpable?

Of course, it can be very hard to tell in any given case whether the agent’s irrational ignorance is badly motivated or not. That is why, in constructing cases for the purposes of inspecting our sense of culpability, we should make it as clear and vivid as possible that the agent’s pattern of concern is entirely appropriate. But some help from the reader cannot hurt: I ask the reader to imagine the characters as being as goodwilled as they could possibly be—as the sort of people who really are concerned with morality and others’ interests as much as they should be, but who suffer from bouts of ordinary unmotivated irrationality.

2. SOME CASES

Sher discusses cases of “poor judgment” for which we are likely to blame the agent.¹⁵ In *Home for the Holidays*, Joilet, a homeowner afraid of burglars, hears movement in the kitchen. She grabs her gun, runs downstairs, and shoots the intruder, only to discover that it was her son who came home early for the holidays. In *Colicky Baby*, Scout gives a crying baby a mixture of vodka and juice to ease its digestive pains.

In each of these cases, it is possible—in fact, easy—to interpret the agent as acting from improper concern. We suspect that Scout knows, at least at some level, that giving a baby vodka is dangerous but just does not care enough. It is also *possible* that the agent is properly concerned but merely foolish, which is the interpretation Sher intends. Perhaps Scout would *never* give a baby something she thought was dangerous, but received some bad advice about how to

14 Kahneman and Tversky, “Judgment Under Uncertainty.”

15 Sher, *Who Knew?* 26.

calm crying babies that she credulously accepted, and now she sincerely believes that giving a baby vodka is normal and safe. Sher contends that these agents are blameworthy, despite their unawareness, and I agree.

Given that it is so tempting to read improper concern into these agents, I think it is helpful to focus on a more detailed case. Consider the following:

Bridge Collapse: Marie is in charge of conducting a stress test on a local bridge, which involves placing weights on the span to see how it responds. Marie has the option of closing the bridge during the test, which is costly and inconvenient, but sometimes required as a precaution. Whether to close the bridge is a matter of judgment reserved for the engineer; the engineer is charged with determining whether the risk of harm is “significant” and then determining whether it should be closed based on this classification. Marie is an experienced and responsible engineer, who always closes the bridge when she believes a stress test is too risky, and usually makes the right call.

But as Marie deliberates on this occasion, she commits a serious error of professional judgment. Given the magnitude of the stress test being conducted and the age of the bridge, it would in fact be *highly* imprudent not to close it. Marie ought to know this, given her training and the evidence (which is extremely clear). She should conclude that the risk of collapse is significant and bring to bear her knowledge that in such cases the bridge should be closed. She is fully capable of drawing this inference; but she does not.

Tragically, the stress test places too much weight on the bridge span and the bridge collapses, violently killing a pedestrian. Marie blames herself for the death, as does the public, once they realize it was the result of her serious error. It is wholly clear to Marie after the collapse that her decision to keep the bridge open was highly irrational, a textbook case of poor judgment on the part of an engineer.

Marie is not improperly concerned. For Marie to have acted from improper concern, her failing to close the bridge must have expressed either indifference to moral considerations or responsiveness to sinister considerations. The relevant moral consideration is the danger that the stress test poses to bridge users. However, Marie is not indifferent to the risk she imposes on others; she always closes the bridge when she believes a stress test is too risky. Her pattern of concern is impeccable in this respect. The problem is that she is unaware, on this occasion, that the stress test is too risky, so she does not know to close the bridge.

Might her ignorance express improper concern? It would if it were motivated

by, say, a desire to avoid having to close the bridge on this occasion (perhaps out of laziness). If it were, Marie's ignorance would express a pattern of improper concern consisting of her overvaluing her own convenience, even at the cost of serious risk to others.¹⁶ But Marie's ignorance was simply the result of a concern-neutral cognitive error, not improper concern of this sort.

If one is still tempted to posit improper concern to explain Marie's error, it may be helpful to think of her ignorance as produced by the operation of a cognitive bias (since this can positively replace an explanation that appeals to insufficient concern). Perhaps it is a bright sunny day and Marie has difficulty imagining the possibility of a disaster on such a lovely day. People tend to underestimate the probability of bad outcomes on sunny days because bad outcomes are less likely to come to mind on sunny days, and thanks to the *availability heuristic*, people often estimate the probability of an event based on the ease with which similar events come to mind.¹⁷ Since Marie's ignorance can be fully explained by the operation of this "cold" bias, there is certainly no need to posit improper concern to explain why Marie neither knows nor suspects the stress test is risky.

Perhaps there is another way Marie could be improperly concerned: maybe Marie just does not care enough about the rationality of her own beliefs. Rosen observes that "we are under an array of standing obligations to inform ourselves about matters relevant to the moral permissibility of our conduct: to look around, to reflect, to seek advice, and so on."¹⁸ We are not *always* morally obligated to check ourselves for bias; I do nothing morally wrong when I complacently believe that Linda is more likely to be a feminist and a bank teller than just a bank teller, or idly question whether the moon landing really took place. But there are cases in which one plausibly should check oneself for irrationality, because the stakes are high and the cost of checking oneself is low. Maybe Marie displays improper concern by not taking more measures to ensure that her judgment is sound, for instance, by seeking advice.

But it is not plausible that Marie displayed improper concern by not seeking advice or rechecking her belief, because, as stipulated, whether to close the bridge is a matter of judgment reserved to the engineer according to the opera-

16 Nor is Marie's ignorance the result of an epistemic vice. FitzPatrick argues that we can be responsible for ignorance that was the result of "voluntary exercise of vices," such as "overconfidence, arrogance, dismissiveness, laziness, dogmatism, incuriosity, self-indulgence, contempt, and so on" ("Moral Responsibility and Normative Ignorance," 605); this is not one of these cases.

17 It is not that the bias renders Marie incapable of getting the right answer; she is a competent engineer and so fully capable of making the right classification. The sunny day makes the problem slightly harder, but not beyond her power by any means.

18 Rosen, "Culpability," 63.

tive procedural norms. The engineer is given latitude to determine whether the risk of harm is “significant,” so Marie is entirely within protocol when she makes the call based on her own deliberation. It is not plausible that Marie displays insufficient concern for others’ safety by failing to go *above and beyond* the protocol, which she has no reason to suspect is flawed. We can even imagine that Marie rechecks her informal calculation, but is subject to the same bias each time, so continues to underestimate the danger posed by the stress test.

There is thus no reason to think that Marie’s pattern of concern is defective (and in any case, we can just stipulate that this is the case). Now the critical question: Are Marie and other comparably irrational agents blameworthy for their foolish decisions? I believe they are. Consider the sort of excuses they might make. Scout in Colicky Baby might sincerely plead that she did not know that vodka was dangerous, that she meant well. Marie might maintain that she followed proper protocol, that she just made an honest mistake. These assertions could all be true, and they justify a certain amount of sympathy for Scout and Marie. But I am inclined to think that they are only partial excuses, even though they establish the goodwill of the agents. Marie’s foolishness *killed* a person. Scout gave a baby alcohol poisoning. These agents were not *incapable* of exercising better judgment. Marie had never made an incorrect call before, and while a sunny day might affect her judgment, nobody can claim she rationally *debilitated* it and that that made her incapable of reasoning correctly. Scout is a normal adult perfectly capable of recognizing that “just give the baby vodka” is poor advice. Our complaint against these agents is simply that they could have and should have known better.

To be clear, my claim is not just that these agents are liable to criticism or that they should compensate the victims if they can. I think they are fully blameworthy, in the sense of meriting “angry blame”: resentment and indignation. To be sure, they merit much less blame than agents who perform comparable acts intentionally. But they are culpable nonetheless.¹⁹

19 It is also worth noting that Anglo American criminal law allows liability in some cases of unreasonable ignorance without requiring any further finding of ill will. A successful mistake-of-fact defense requires that the defendant demonstrate his mistake was reasonable: one that an ordinary person would (or might) have made under the circumstances. In *Commonwealth v. Pierce* (138 Mass. 165 [1884]), a leading case on ignorance of fact, the defendant was a doctor who attempted to cure a patient by wrapping her in kerosene-soaked clothes. Though the doctor was determined to be acting without knowledge of the risk, he was convicted of involuntary manslaughter on the ground that his ignorance was unreasonable. Such agents may be less culpable than people who cause harm knowingly or recklessly. But it is simply not part of law to excuse them altogether on the ground that they were irrational or stupid or otherwise unreasonable rather than ill willed.

One might worry that, despite every stipulation to the contrary, our sense of culpability is influenced by an inclination to attribute ill will or “mismanagement of opinion” to these agents. Rosen argues that though we are sometimes inclined to blame agents who act wrongly from blameless ignorance, this is due to a failure to hold clearly in mind all the relevant facts. When we *do* manage to bear in mind that the agent’s ignorance was blameless, our resentment evaporates, as Rosen claims in this discussion of our reaction to an ignorant ancient slaveholder (who is unaware that slavery is morally wrong):

We are no doubt powerfully inclined to blame [the ancient slaveholder]—so long as we ignore the stipulated fact that he is blameless for not knowing that his slave deserves much more. When we bear this in mind—when we “zoom out,” as it were—then (I claim) our sense of his culpability evaporates. It is as if in blaming him we are thinking that he should have known better; he should have known that his action expresses an attitude that would merit intense resentment; he should have known that his act is the sort of act for which one should feel shame in retrospect. When we remind ourselves what it would have taken for him to know these things, and in particular that it’s not his fault that he doesn’t know them, then our resentment is properly blocked.²⁰

So, for instance, if we manage to hold in mind that Marie followed proper protocol—and hence, her ignorance was blameless—our sense of culpability should evaporate according to Rosen. But I do not find Rosen’s claim persuasive here. Rosen claims that the “should” thoughts that sustain resentment are falsified when an agent’s ignorance is blameless: “he should have known better,” “he should have known that his action expresses an attitude that would merit intense resentment,” and so on. These thoughts are indeed false if the “should” is interpreted (as it is by Rosen) as indicating the agent should have *done* something differently in the management of her opinion.²¹ But this is not the only way to interpret the relevant “should” thoughts that trigger resentment. If the agent’s ignorance is unreasonable, then a different, epistemic version of the “should” thought is available: “*X* *rationally* should have known better.” Her epistemic performance fell below some operative epistemic standard. I maintain that

20 Rosen, “Culpability,” 73.

21 According to Rosen, someone is blameworthy in her ignorance only if it can be traced to some prior blameworthy act or omission, like a failure to satisfy one’s procedural-epistemic obligations. Rosen stipulates that the slaveholder complied with his procedural-epistemic obligations and hence his ignorance is blameless; “we are under no obligation to rethink the uncontroversial normative principles that form the framework for social life” (“Culpability,” 65).

thoughts about the agent's epistemic failings suffice to sustain resentment and indignation in these cases. These emotions are stoked simply by the thought that the agent could have and should have (according to some operative epistemic standard) known better, and this is so even if we expressly cancel the thought that the agent's failure of rationality is fueled by ill will on her part. We imagine the agent having all the evidence in front of her, in full possession of her cognitive faculties, yet drawing an utterly unreasonable conclusion, to dire effect. This is enough to fuel resentment's fire.

Of course, just because we *tend* to resent agents who exercise poor judgment does not mean we are *correct* to resent them. One might dismiss this inclination as seriously misguided, and call for major revision to our ordinary practice of blame and punishment. But I do want to point out that this is in some tension with the Strawsonian approach to moral responsibility that originally motivated the ill-will condition. Strawson believed that any credible account of the conditions of responsibility must vindicate ordinary practice at least to a significant degree: our practice of holding one another responsible is an "essential feature of our way of life" and therefore "neither calls for nor permits, an external 'rational' justification."²² Strawson thought ill will was necessary for blameworthiness because this is what ordinary practice reflects: "The personal reactive attitudes rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves."²³ But I have claimed that ordinary practice with the reactive emotions does not always reflect a demand for goodwill: it sometimes reflects a demand for *reasonableness*, as when we resent agents for their poor judgment. If that is right, then Strawsonians who consider the ill-will condition to be vindicated by the standards of ordinary practice have a predicament.

To see why Strawsonians of this sort are in trouble, consider Rosen's Strawsonian view. Rosen is a Strawsonian in the sense that he both (1) endorses the ill-will condition and (2) believes this is vindicated by an aspect of ordinary practice, "what the personal reactive attitudes rest on."²⁴ According to Rosen, resentment does not only *depend* on the belief that the agent expressed ill will, but *constitutively involves* this thought. That is, just like fear of *X* contains the thought "*X* is dangerous," resenting *X* for *A* contains the thought that "*X*'s doing *A* expressed an objectionable pattern of concern."²⁵ And this is why, for Rosen, agents are blameworthy when they express ill will: expressing ill will makes or-

22 Strawson, *Freedom and Resentment and Other Essays*, 23.

23 Strawson, *Freedom and Resentment and Other Essays*, 17.

24 Strawson, *Freedom and Resentment and Other Essays*, 347.

25 Rosen, "The Alethic Conception of Moral Responsibility," 14.

dinary resentment appropriate, given its constituent thought. (Just like fear of X is appropriate just in case X is *in fact* dangerous.)

But is this right? On a view like Rosen's, we can get a handle on the thoughts implicit in an emotion by asking which sort of discoveries cause it to evaporate. For instance, if fear usually evaporates upon learning or perceiving that X is not dangerous, this is evidence that "X is dangerous" is a constituent thought of fear. Rosen contends that resentment always evaporates when we "zoom out" and bear in mind the agent's ignorance was blameless. But I have argued that the thought "X could have and should have known better" is enough to sustain resentment, even in the absence of ill will. If that is right, this suggests the ill-will thought is not constitutive of resentment.²⁶ And this is trouble for the foundation of Rosen's account, since it removes the very justification for the ill-will condition.

I concede that these arguments are not decisive. Whether we regard unreasonable ignorance as sufficient for blameworthiness should ultimately depend on the overall plausibility of a theory that rejects the quality of will constraint and provides alternative conditions of blameworthiness. My aim in what follows is to sketch the beginnings of such a theory.

3. RECOVERING THE VIEW OF BLAMEWORTHINESS

Ill-will theorists have a pleasingly simple theory of blameworthiness: X is blameworthy for a wrong act A iff in doing A , X expressed ill will (and, perhaps, possessed a certain requisite capacity, like the capacity to assess and respond to moral reasons—what Strawson called a "moral sense").²⁷ This theory must be revised if we are to account for the culpability of unreasonable but non-ill-willed

26 One might object to this argument on grounds that it is clearly possible to experience a reactive emotion inappropriately. Fear contains an "X is dangerous" thought, but someone can recoil in fear at the sight of what she knows to be a harmless snake. Might resentment for unreasonableness be like this—possible and understandable but not appropriate? It is worth noting that even when one knows that a snake is harmless, it can still give the *appearance* of being dangerous—we are hardwired to see the serpentine silhouette as indicating "danger." Hence one might have a quasi-perceptual "X is dangerous" thought, in spite of one's explicit knowledge. But we resent for unreasonableness even when it is *manifestly clear* that the agent was not ill willed (as in Bridge Collapse). Though it is always *possible* that our resentment is based on some illusion of ill will, the best explanation for this, I believe, is simply that we can resent without thinking anything like the ill-will thought. As Rosen acknowledges: "An account of the thoughts implicit in resentment is unacceptable if it is clear on reflection that we can resent X for A without thinking the thoughts posited by the account" ("The Alethic Conception of Moral Responsibility," 75).

27 Strawson, *Freedom and Resentment and Other Essays*.

agents. How might we do so? One possibility is to resign ourselves to an unappealing disjunctive account: X is blameworthy for a wrong act A iff in doing A , X expressed objectionable will, *or* unreasonable ignorance (perhaps, again, against the backdrop of possessing the capacity to assess and respond to reasons). This would be to concede that sometimes we resent agents in virtue of ill will, and sometimes unreasonable ignorance, and that no more unified account of our practice can be given.

But such resignation is unnecessary. There is, I believe, a unified view of moral blameworthiness consistent with our practice that accounts for the blameworthiness of agents who act from unreasonable ignorance, and ill will (though the blameworthiness of ill-willed agents will not lie in their ill will, *per se*, but in their unreasonableness). On this account, agents are blameworthy whenever they act wrongly from an unreasonable attitude—a belief, intention, or desire—that is attributable to them, in a sense I will explain. Acting from unreasonable ignorance and acting from ill will are both ways of acting from unreasonable attitudes attributable to the agent.

To see how cases of blameworthy ill will and blameworthy ignorance might be unified, let us observe that, in both sorts of cases, the agent's wrong act reveals an unreasonable attitude (belief, intention, or desire): one that is held on the basis of bad reasons, or is insufficiently sensitive to the reasons the agent has. This assumes that there are rational norms on final desires as well as beliefs, something that Humeans have long denied. But let us suppose that final desires are rationally criticizable, so that, for instance, a person who finally desires to inflict suffering on others wants something she has no good reason to want. Then wrong action from bad desire and wrong action from irrational ignorance both express unreasonableness, as can be seen in the following pair of cases:

Ill Will: Suppose Smith kills Jones out of malice. There are conclusive reasons for Smith not to kill Jones (just as there are conclusive reasons for anyone not to kill gratuitously), but Smith is irrationally insensitive to these reasons. Hence the *intention* or *desire* that moves Smith to act is unreasonable.

Unreasonable Ignorance: Suppose Smith shoots and kills Jones while horsing around due to a false belief that playing around with a gun is not risky. Smith has conclusive reason to believe that horsing around with a gun is risky, but is irrationally insensitive to these reasons. Hence the *belief* from which Smith acts is unreasonable.

Of course, the sort of reasons the agent improperly assesses in each case are dif-

ferent. In Ill Will, the relevant reasons are reasons for intending or desiring. In Unreasonable Ignorance, the relevant reasons are reasons for belief. But insofar as each is subject to rational norms, both agents act from attitudes that are rationally criticizable.

To be sure, in Ill Will Smith also expresses ill will in the ordinary sense; he has a final desire to do what is in fact wrong. A traditional Strawsonian will say that this bad desire suffices to make him blameworthy, regardless of whether it is rationally faulty. What I am considering is that it is not the badness of the desire that renders Smith culpable, but its unreasonableness.

Michael Smith argues for something like this view: it is irrationality in either an agent's beliefs or desires that renders her "at fault" for wrong conduct.²⁸ First, consider Gary Watson's contrast between self-indulgence, weakness, and compulsion, to which Smith appeals:

Suppose that a particular woman intentionally takes a drink. To provide an evaluative context, suppose she ought not to have another because she will then be unfit to fulfill some of her obligations. Preanalytically, most of us would insist on the possibility and significance of the following three descriptions of the case: (1) the reckless or self-indulgent case; (2) the weak case; and (3) the compulsive case. In (1), the woman knows what she is doing but accepts the consequences. Her choice is to get drunk or risk getting drunk. She acts in accordance with her judgment. In (2) the woman knowingly takes the drink contrary to her (conscious) better judgment; the explanation for this lack of self-control is that she is weak-willed. In (3), she knowingly takes the drink contrary to her better judgment, but she is the victim of a compulsive (irresistible) desire to drink.²⁹

Smith contends that we blame both the self-indulgent woman and the reckless woman. We blame the self-indulgent woman for "having the wrong belief about what she should do in the circumstance" when she could and should have known better. We blame the weak woman, "[not] for her belief—she has the belief she should have, after all—but rather for her failure to act on that belief" when she could have.³⁰ By contrast, we do not blame the woman who acted compulsively at all, because she believed correctly and could not have done otherwise than she did.

Having the wrong belief (like the self-indulgent woman) and having the wrong intention or desire (like the weak woman) are both forms of irrational-

28 Smith "Rational Capacities."

29 Watson, "Skepticism about Weakness of Will," 324.

30 Smith, "Rational Capacities," 18.

ity. The self-indulgent woman fails to recognize the reasons against intending/desiring to take the drink; the weak woman recognizes these reasons, but fails to desire in accord with them. Says Smith:

If people possess, but fail to exercise the relevant [rational] capacities, then there is something that they could have done in a perfectly mundane sense: they could have exercised the capacity to access the available evidence, or to believe in accordance with their evidence, or to . . . desire, in the way that's rationally required. Their failure to access or believe or . . . desire correctly is therefore their fault . . . and so too is their failure to perform the act that they would otherwise have performed.³¹

As Smith is careful to emphasize, agents must have the *capacity* to assess reasons, and to modulate their beliefs and desires accordingly, in order to be held responsible for their unreasonable conduct. Nearly everyone agrees that very young children, animals, and the psychotic are not liable to blame even when their acts are sourced in unreasonable intentions, desires, or beliefs. This is presumably because they substantially lack the capacity to assess reasons and to believe and desire in line with assessments of reasons, so their failures to do so are not their fault.

In saying their unreasonableness is not their “fault,” we are not yet making a point about moral blameworthiness or the lack thereof in the sense at issue in this paper. As I read him, Smith’s talk of “fault” in this context concerns a different notion: whether the failure is properly attributable to the agent, i.e., whether she is properly criticized *qua* agent on the basis of it. If an agent possesses the relevant capacity, then—assuming the capacity is not masked (because she has been drugged or is excessively tired or whatnot)—failures to properly exercise the capacity are “down to her.” But if she lacks the capacity, they are not “down to her” (they are, rather, attributable to the absence of the capacity), and she is not criticizable for these rational failings.

So, having the relevant capacity—that is, the capacity to assess reasons and modulate one’s attitudes in light of these assessments—renders one responsible, *in the attributability sense*, for one’s attitudes. That is Smith’s point to which I add: agents are blameworthy *in the resentment-involving sense* when they act from unreasonable attitudes that are attributable to them.

Taking account of all this, the resultant theory of blameworthiness, which I will call *Rational Capacitarianism*, is as follows:

31 Smith, “A Constitutivist Theory of Reasons,” 11.

Conditions of Blameworthiness: X is blameworthy for a wrong act A iff X 's doing A expressed an unreasonable attitude that is attributable to her.³²

Theory of Attributability: An attitude is attributable to X iff X has the capacity to assess the reasons for and against the attitude and modulate her attitude in light of this assessment during the period in which the attitude was formed and sustained.³³

Take again a case of blameworthy ill will: Smith kills Jones out of malice. On Rational Capacitarianism, Smith is blameworthy because his wrongful act is rooted in some unreasonableness that is down to him:

1. What is the unreasonable attitude expressed? The intention or desire to kill Jones.
 2. Why is the intention attributable to Smith? Smith is a competent adult and had the ability to assess reasons for and against intending/desiring, and to modulate his intentions/desires accordingly when the attitude was formed and/or sustained.
 3. Why is the intention unreasonable? There are two possibilities. If Smith fails to recognize that there is conclusive reason not to kill Jones despite awareness of the pertinent facts (analogous to the self-indulgent woman), his unreasonableness lies in his flawed belief. If he recog-
- 32 The "expression" of an irrational attitude, we might say, involves its appearance in a certain standard story of action. On the standard story, an agent's ϕ -ing is an action when it results non-deviantly from her beliefs and desires combining in a certain way, i.e., when she (1) has some final desire to ψ and (2) a belief that ϕ -ing is likely to bring about her ψ -ing, which results in an instrumental desire to ϕ . An action "expresses" an irrational belief or desire when one of the beliefs or desires that produces the action in this way is irrational.
- 33 My account bears similarity to Fischer and Ravizza's reasons-responsiveness account of moral responsibility, but should be distinguished. Fischer and Ravizza provide a necessary condition for responsibility: a condition that an agent must satisfy if she is to be morally responsible for her actions. By "responsible," they simply mean an apt target of the reactive attitudes, both negative and positive: "Our Strawsonian view of moral responsibility allows for responsibility for 'morally neutral' behavior. For instance, one can be morally responsible for simply raising one's hand" (*Responsibility and Control*, 8). In contrast, I am arguing that failures of the capacity for reasons-responsiveness (in the sense I have described) are sufficient for blameworthiness when expressed in wrongful action, and moreover, failures of the capacity ground an agent's blameworthiness in cases of wrongful action from ignorance. (Fischer and Ravizza do not tailor their account to ignorance; their focus is "the freedom-relevant condition on moral responsibility, rather than the epistemic condition"; *Responsibility and Control*, 13.) Someone could accept Fischer and Ravizza's view that reasons-responsiveness is a necessary condition for responsibility without endorsing my view that failures of this capacity are sufficient for, and moreover are the ground of, blameworthiness when expressed in wrongful action.

nizes there is conclusive reason not to kill Jones but kills him anyway (analogous to the weak woman), his unreasonableness lies in his flawed intention or desire.

Now return to Marie in Bridge Collapse. Marie also acts from an unreasonable attitude that is attributable to her: the irrational belief that the stress test is not risky. Why is the attitude unreasonable? It reflects an improper assessment of the considerations that favor closing the bridge. Why is it attributable to Marie? Marie could have evaluated the evidence properly and believed in accord with it; she was not rationally debilitated. Hence the collapse is Marie's fault, and she is blameworthy for it.

It is not obvious that what makes an attitude attributable to an agent, of course, is that it expresses the failure to exercise a rational capacity of this sort, though this is what bad desires and unreasonable ignorance appear to have in common. One might alternatively maintain that an attitude is attributable to an agent only if it actually reflects her evaluative judgments. This is the view of Angela Smith:

To say that an agent is morally responsible for something ... is to say that that thing reflects her rational judgment in a way that makes it appropriate, in principle, to ask her to defend or justify it.... Most of our desires, beliefs, and other attitudes seem to meet this condition of judgment-dependence, even though they do not commonly reflect a choice or decision, and are not normally under our voluntary control.³⁴

The problem with this view is that it cannot explain responsibility in cases of weakness of will. Both the self-indulgent and the weak woman are responsible in Watson's example, but only the self-indulgent woman's act reflects her evaluative judgment. The weak woman acts *against* her better judgment, taking the drink when she knows she should not.

What unifies the class of attitudes that are attributable to us, I believe, is not that they actually reflect our evaluative judgments, but merely that they are *apt* to reflect them—they are the sorts of attitudes that *could* be brought into alignment with our evaluative judgments. These attitudes, which include beliefs, desires, and intentions, are under a kind of "rational control," in the sense that they are capable of being regulated by, hence controlled by, an agent's recognition of norms, however imperfectly.³⁵ Some writers see this sort of "control by the

34 Smith, "Control, Responsibility, and Moral Assessment," 369–70.

35 The presence of "rational control" is what leads Pettit and Smith to remark that "[An agent's] beliefs do not just come and go in a natural procession of events.... The subject is

agent's view of the reasons" as strongly analogous to voluntary control, and so as amounting to a kind of freedom akin to free will.³⁶ But that is not necessary for present purposes. All that matters is that when an attitude is formed in an agent who possesses the capacity to respond appropriately to reasons and to modulate his attitudes in light of his assessment of the reasons, that attitude is, in every morally relevant sense, his own.

4. CAPACITARIANISM

There is another account of responsibility that widens the scope of attributable attitudes. On this view, not only failures of rational capacities are attributable to the agent, but failures of any capacity, or perhaps of any "psychological capacity." Capacitarians agree with me that quality of will is not necessary for blameworthiness.³⁷ Capacitarians contend that agents with a decent quality of will are blameworthy for wrong acts performed from ignorance if, roughly, they could have and should have known that the act had the relevant features—that is, if their ignorance is the result of the improper exercise of some capacity.³⁸

Capacitarians tend to take cases of forgetful negligence as paradigm cases of culpable ignorance, which I have deliberately avoided doing. For instance, Sher focuses on a case that he calls *Hot Dog*, in which Alessandra forgets her dog Sheba in a hot car. Alessandra forgets about Sheba not because she cares too little about Sheba, but because she gets distracted by a situation occurring at her child's school. Sher contends that Alessandra is blameworthy because she

certainly not a mere passive or mechanical system. She does not just revise her beliefs and desires autonomically, or at any rate, not when they operate beyond the reach of the occasional disabling obstacles that get in her way. She revises them under the spur of recognizing what the relevant norms require of her" (*Freedom in Belief and Desire*, 442). Some have argued that a kind of capacity-relative control suffices for an agent's ignorance to itself be blameworthy (Rudy-Hiller, "A Capacitarian Account of Culpable Ignorance"). For instance, on Rudy-Hiller's capacitarian account, "an agent can be directly in control (in the responsibility-relevant sense) of his ignorance, since he can have direct capacitarian control over whether he notices, remembers, or otherwise is aware of a relevant consideration" (415). And this means an agent can be "directly responsible" for his ignorance (in the sense of being culpable for it) (407). *Contra* Rudy-Hiller, I think agents can be blameworthy for these attitudes themselves; an agent's defective attitude must be expressed in volitional action to be blameworthy.

36 See, e.g., Smith and Pettit, "Freedom in Belief and Desire."

37 Sher, *Who Knew?*; Rudy-Hiller, "A Capacitarian Account of Culpable Ignorance"; Clarke, "Negligent Action and Unwitting Omission"; Murray, "Responsibility and Vigilance"; and Amaya and Doris, "No Excuses."

38 Rudy-Hiller, "A Capacitarian Account of Culpable Ignorance," 405.

could have and should have remembered that Sheba was in the car; her failure of memory fell below “some applicable” standard that she could have met.³⁹

But we must be careful, as we are not always responsible for failing to exercise capacities we possess. Consider the following case.

Cheerleading Disaster: Sarah is a base on the cheerleading squad tasked with catching Jane, a flyer, after an air flip. Sarah is strong enough to catch Jane and has done so many times before, so she has the capacity to catch her. But on this occasion, her muscles fail her, and Jane falls through her grip, resulting in serious injury to Sarah.

The problem is not a failure of intention or effort on Sarah’s part; Sarah exerted more than enough effort as would normally be needed. Her muscles simply did not cooperate with her intention. It is true that Sarah “could have and should have” caught Jane, but this is intuitively insufficient for her moral blameworthiness when we are clear about why she did not do what she could have and should have done. The challenge for capacitarrians is to distinguish morally relevant capacities from irrelevant ones.⁴⁰

Rational capacitarrians draw the line at rational capacities; all and only failures of rational capacities (including the capacity to form evaluative judgments and modulate one’s attitudes accordingly) are attributable to the agent. The underlying idea is that an agent is responsible only for her attitudes that are apt to reflect her evaluative judgments, that *could* be brought into alignment with them, and would if we were perfectly rational—the attitudes that are under “rational control.”

This raises the question: Is what we remember under rational control? Are failures of memory attributable to us; are we answerable for them? This is a difficult question. To be sure, what we remember is sensitive to some degree to our judgments about reasons. Suppose at 8 AM I think to myself that I really ought to pick up some groceries after work. This makes me more likely to remember to pick up groceries after work. If one judges something to be very important, this increases the chance that thoughts about it will come to mind at the right time, that one’s memory will be jogged by things pertaining to it, and so on. If a failure of memory is traceable to a failure to recognize the strength of the reasons

39 Sher, *Who Knew?* 88.

40 Sher contends that for an agent to be responsible for a failure of a capacity the failure must be “caused by the interaction of some combination of his constitutive attitudes, dispositions, and traits” (*Who Knew?* 88). I am skeptical that Sher’s proposal can explain all the cases. Marie’s failure to draw the correct conclusion in Bridge Collapse is the result of her being subject to certain biases; and being subject to these does not seem constitutive of who she is.

for carrying out the task, it is traceable to a failure of a rational capacity and the agent is responsible for it.

The harder case is when an agent recognizes the importance of performing the task, resolves to do it, and then simply forgets to bring about the planned action. This sort of mnemonic failure seems relevantly different from a failure to intend or believe in accordance with one's reasons. With belief and intention, rationality concerns roughly the synchronic harmony between one's assessments of reasons and one's first-order attitudes: ideally rational agents bring them into synchronic consistency. When a rational agent is aware of a reason, forms a judgment about what reason requires, and adjusts her first-order attitudes so as to bring them into conformity with this judgment, the episode, though not strictly voluntary, does seem to involve a kind of activity and is properly read as the agent's making the resulting attitudes her own.

Remembering, on the other hand, does not seem to involve a kind of rational activity. Resolving to remember seems more like setting an alarm clock and hoping it goes off at the right time; the episode of recollection itself either happens or not, though not in a way traceable to the self. Failures of memory seem more like "glitches" analogous to the failure of strength in *Cheerleading Disaster* than defects in the agent. Consider that, in forgetting to perform some task, one's beliefs about the reasons to perform the task are not occurrent, so remembering is not merely a matter of bringing one's assessments of reasons and one's first-order attitudes into synchronic consistency. It seems to me, therefore, that a failure of memory of this sort is not a failure of a rational capacity in the relevant sense, so is not attributable to the agent.

This may be controversial; to the extent that memory is under rational control in the sense in which beliefs, desires, and intentions are, it may make sense to blame agents for failures of memory. But precisely because this is unclear, I think it is a mistake for ill-will skeptics to emphasize forgetting cases. When we focus on cases of glitchy memory in goodwilled agents and stipulate beyond all doubt that at every stage along the way they responded appropriately to the reasons of which they were aware, exercising impeccable rational control, it is hard to see the agent as anything but the passive victim of a glitchy sub-personal module on which he was fully entitled to rely. Cases of this sort are at any rate certainly not clear cases of responsibility. Cases of rational failure on the part of competent agents, I believe, are much clearer. This theory is, of course, only a sketch; but it shows how blame for "ill will" may be more continuous with blame for unreasonable ignorance than one might have initially supposed. I hope to have shown that we do not need to choose between a quality of will view (which seems too lenient) or a standard capacitarian view (which seems too harsh). Ra-

tional capacitarianism—on which only rational capacities are relevant—offers a principled alternative.

5. CONCLUSION

I have argued that agents are blameworthy when they act wrongly from unreasonable ignorance even in the absence of ill will, and provided a theory that ratifies this verdict. The theory entails that when an agent acts wrongly from factual ignorance, she is blameworthy only if her ignorance derives from a failure to exercise a rational capacity—a capacity to recognize reasons for believing and desiring and to be responsive in one's beliefs, desires, or intentions to one's assessment of reasons.

The plausibility of a theory that licenses blame for non-ill-willed unreasonableness in part depends on what blame and blameworthiness consists in. In closing I will say a few words about this issue. Like Rosen, Gibbard, and Wallace, I have fastened on a conception of blame according to which the primary forms of blame are the negative reactive emotions like resentment and indignation, and on which an agent is blameworthy for an act iff she is liable to resentment and indignation for it.⁴¹ But this is not a complete analysis, because it does not specify what being “liable” to the negative reactive attitudes amounts to. How we fill this gap is important.

Rosen's view, for instance, is that since resentment and indignation are emotions their appropriateness conditions are given by their constitutive thoughts. Emotions involve belief-like mental states; fear involves the thought that *X* is dangerous, misery the thought that things are going badly, and so on, and an emotion is appropriate just in case its constituent thoughts are true.⁴² Accordingly, for it to be appropriate to resent *X* for *A* is for the thoughts implicit in resentment to be true of *X* and *A*.⁴³

Everything I have said in this paper is consistent with this general view. If we adopt this framework, it should be read as contesting Rosen's view of the constitutive thoughts of resentment. Rosen thinks there is a “quality of will thought” contained in resentment (“in doing *A*, *X* expressed an objectionable pattern of concern”), and I disagree. My argument for this is simply that it is possible to resent *X* for an act that expresses unreasonable ignorance, but not ill will, clear-

41 Rosen, “The Alethic Conception of Moral Responsibility”; Gibbard, *Wise Choices, Apt Feelings*; and Wallace, *Responsibility and the Moral Sentiments*.

42 Rosen, “The Alethic Conception of Moral Responsibility,” 71–72.

43 Rosen, “The Alethic Conception of Moral Responsibility,” 72.

headedly, in full awareness of this fact. Resentment is stoked in these cases by the thought that *X* could have and (rationally) should have known better.⁴⁴

If we reject Rosen's account of what it is to be "liable" to the reactive emotions, it may be harder to see how resentment directed at unreasonableness could be appropriate in the absence of ill will. The "fairness view" specifies that *X* is liable to resentment for *A* just in case it would be morally fair to resent her for *A*, and the "fittingness view" appeals to a primitive, *sui generis* relation of fittingness.⁴⁵ But how could it be fair, or fitting, to resent—which involves the withdrawal of some amount of goodwill—if there was no comparable ill will on the part of the target of the resentment?

I do not have a complete answer to this question, but I will outline the beginnings of a response. The fairness of resentment plausibly depends on whether the target of the resentment had a fair opportunity to avoid being a target of resentment. Rudy-Hiller contends, for instance, that blaming for ignorance is appropriate because an agent "can have direct capacitarian control over whether he notices, remembers, or otherwise is aware of a relevant consideration," which is simply a matter of having the "requisite abilities to do these things and a fair opportunity to exercise them."⁴⁶ I differ from Rudy-Hiller in thinking that resentment is only appropriately directed at attitudes that are under the agent's *rational* control, but the spirit of my view is the same. Agents can avoid being the target of resentment by exercising their capacity to recognize reasons for believing and desiring and to be responsive in their beliefs and desires to their assessment of reasons. This is enough, it seems, for a fair opportunity to avoid.

Resentment plausibly is a sort of sanction directed at agents who we think could have done better, aimed at ensuring the agent will do better in the future. The attitudes with respect to which this corrective activity makes sense are precisely the ones that are under the agent's rational control—the attitudes that she can modulate according to her evaluative judgments. What resentment "wants," so to speak, is for the agent to recognize and revise her defective attitude (for instance, by adopting the appropriate belief or desire); this is why resentment is often not satisfied unless the agent recognizes *why* what she did was wrong. If resentment is a correctional activity of this sort, it is easier to see how it could be appropriately directed at all forms of unreasonableness, not just ill will.

44 As Rosen acknowledges, "An account of the thoughts implicit in resentment is unacceptable if it is clear on reflection that we can resent *X* for *A* without thinking the thoughts posited by the account" ("The Alethic Conception of Moral Responsibility," 75).

45 Wallace, *Responsibility and the Moral Sentiments*; see Rosen, "The Alethic Conception of Moral Responsibility," 70.

46 Rudy-Hiller, "A Capacitarian Account of Culpable Ignorance," 415, 407.

I will conclude by stressing that the justification of serious punishment is another matter. Resentment and indignation are moderate, temporary responses that tend to lessen when the agent corrects the attitudes toward which resentment is directed. Serious punishment (e.g., a restriction of an agent's liberty) of the sort that conveys serious moral condemnation is plausibly inappropriate unless the agent expressed improper concern; as the costliness to the agent of a blaming response increases, the more it seems reasonable to require some morally objectionable feature of the agent's will to be involved.⁴⁷

Princeton University
 aayars@princeton.edu

REFERENCES

- Amaya, Santiago, and John M. Doris. "No Excuses: Performance Mistakes in Morality." In *Handbook of Neuroethics*, edited by Jens Clausen and Neil Levy, 253–72. New York: Springer, 2015.
- Arpaly, Nomy. *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford: Oxford University Press, 2002.
- Baron, Marcia. "Culpability, Excuse, and the 'Ill Will' Condition." *Aristotelian Society Supplementary Volume* 88, no. 1 (June 2014): 91–109.
- Clarke, Randolph. "Negligent Action and Unwitting Omission." *Surrounding Free Will: Philosophy, Psychology, Neuroscience*, edited by Alfred R. Mele, 298–317. Oxford: Oxford University Press, 2014.
- Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press, 2000.
- FitzPatrick, William J. "Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge." *Ethics* 118, no. 4 (July 2008): 589–613.
- Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press, 1992.
- Graham, Peter A. "A Sketch of a Theory of Moral Blameworthiness." *Philosophy and Phenomenological Research* 88, no. 2 (March 2014): 388–409.
- Harman, Elizabeth. "Does Moral Ignorance Exculpate?" *Ratio* 24, no. 4 (December 2011): 443–68.

⁴⁷ Thanks to an anonymous reviewer for this point. For invaluable comments on and discussion about this article I would like to thank Gideon Rosen, Randolph Clarke, Michael Smith, Elizabeth Harman, and three anonymous reviewers for the *Journal of Ethics and Social Philosophy*.

- McKenna, Michael. "Directed Blame and Conversation." *Blame: Its Nature and Norms*, edited by D. Justin Coates and Neal A. Tognazzini, 119–40. Oxford: Oxford University Press, 2012.
- Murray, Samuel. "Responsibility and Vigilance." *Philosophical Studies* 174, no. 2 (2017): 507–27.
- Pettit, Philip, and Michael Smith. "Freedom in Belief and Desire." *Journal of Philosophy* 93, no. 9 (September 1996): 429–49.
- Rosen, Gideon. "The Alethic Conception of Moral Responsibility." *The Nature of Moral Responsibility: New Essays*, edited by Randolph Clarke, Michael McKenna, and Angela M. Smith, 65–88. Oxford: Oxford University Press, 2015.
- . "Culpability and Ignorance." *Proceedings of the Aristotelian Society Volume* 103, no. 1 (June 2003): 61–84.
- Rudy-Hiller, Fernando. "A Capacitarian Account of Culpable Ignorance." *Pacific Philosophical Quarterly* 98, no. s1 (December 2017): 398–426.
- Sher, George. *Who Knew? Responsibility without Awareness*. Oxford: Oxford University Press, 2009.
- Smith, Angela M. "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 138, no. 3 (April 2008): 367–92.
- Smith, Michael. "A Constitutivist Theory of Reasons: Its Promise and Parts." *Law, Ethics and Philosophy* 1 (2013): 9–30.
- . "Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion." In *Weakness of Will and Practical Irrationality*, edited by Sarah Stroud and Christine Tappolet, 17–38. Oxford: Oxford University Press, 2003.
- Strawson, P. F. *Freedom and Resentment and Other Essays*. Abingdon, UK: Routledge, 2008.
- Tversky, Amos, and Daniel Kahneman. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157 (September 27, 1974): 1124–31.
- Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1994.
- Watson, Gary. "Skepticism about Weakness of Will." *Philosophical Review* 86, no. 3 (July 1977): 316–39.
- Zimmerman, Michael J. "Moral Responsibility and Ignorance." *Ethics* 107, no. 3 (April 1997): 410–26.