**Accounting for Failure**

Randolph Clarke

Our agency is subject to failure of various kinds. We miscalculate, misjudge, mistake, miss our marks, overlook, forget, slip, trip, drop, or spill. Some of these are failures to do things that are easy for the agents to do, given their capabilities and circumstances. My focus lies within this narrower range of cases.

This group is itself quite varied, the cases differing with respect to whether the failures stem from arrogance, overconfidence, laziness, disregard, inattention, poor reasoning, memory lapse, or something else. In which kinds of case can one be responsible for, and to blame for, one’s failure?

With respect to blame, my concern is with specifically *moral* blame. One might figuratively kick oneself for failing to seize an opportunity, and doing so is a way of blaming oneself. But the failure need not involve moral fault, and any blame of which one is worthy need not be moral blame. Further, I am concerned with a kind of moral blame that can include emotional responses such as guilt, anger, and resentment. Worthiness of such blame is commonly the fittingness of such responses. Moral responsibility that includes blameworthiness of this kind is called *accountability*: one who is blameworthy in this way can properly be held to account for what she has done (though *what* may be done, and *who* may do it, varies greatly from case to case).

Cases on which I focus pose a challenge to a widely held view of moral responsibility, one on which an agent is morally blameworthy for conduct only if it expresses, or is appropriately explained by, objectionable quality of will.[[1]](#footnote-2) The latter is commonly understood to consist in desiring, valuing, or caring about something too little, or desiring, valuing, or caring about another thing too much, or having an objectionable pattern or constellation of desires, values, or cares. Although failures of agency are often due to such objectionable will, often they are not, and many in which objectionable will does not figure are nevertheless, I think, cases in which moral blame is fitting. Reflection on a variety of cases supports a view of responsibility that does not include a quality-of-will requirement for blameworthiness. How to state an alternative is a question to which I offer only a tentative suggestion.

1. Varieties of Failure

Failure can result from poor judgment even when the agent is generally a good judge of the relevant matters. Consider **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…. 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 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. (Ayars 2021: 62)

Marie, it can be stipulated, is well motivated, and she is under no special pressure to keep the bridge open, nor unduly distracted by other concerns. She does not have epistemic vices of over-confidence, arrogance, dismissiveness, laziness, incuriosity, etc.[[2]](#footnote-3) Her decision is not reckless; she doesn’t judge it risky to keep the bridge open but recommend doing so anyway; rather, she misjudges the risk. Such mistakes are made even by those who are in general good judges of the relevant matters. Still, the bridge collapse is her fault, and the family of the victim may fittingly blame her. If they subsequently forgive, they would not be making the mistake of forgiving someone who is blameless.

Poor judgment can itself result from failure to take into account information that one possesses but fails to consider in deliberation. Consider **Bad Dog**:

Bob is watching TV when his eight-year-old son walks up. The boy often accompanies Bob on walks with their dog, and he asks to take the dog out himself. Pausing the show, Bob asks where and, satisfied with the answer (and pleased to see the boy take initiative), he reminds his son to pick up after the dog. During the walk, the dog lunges aggressively at a neighbor’s dog and escapes the boy’s grasp. He badly bites both the neighbor and her dog. He had behaved aggressively toward other dogs [on previous occasions], but Bob hadn’t thought of this fact when he considered the boy’s request. (Clarke 2017: 239)

Bob should have judged that it was too risky to send his son out alone with the dog. He possessed information that clearly favored that assessment. But he failed to think of this consideration while deliberating—his attention was drawn to the pleasant fact that his son had asked to do a chore—and consequently Bob misjudged the matter. He has, we may imagine, normal capacities for practical deliberation, including capacities to employ relevant information that he possesses, even in contexts of such pleasant facts. Bob failed to exercise his capacities well on this occasion. The outcome was bad, and it is Bob’s fault. It would be fitting for him to feel guilty for his mistake. Angry blame by the victim of his mistake would not be out of place.

Some failures result from failures to perceive, which can in turn stem from failures to attend sufficiently. Here is an example, **Stop Sign**:

Ann is driving to a friend’s house when she collides in an intersection with another car, killing one of its passengers. Ann has run a stop sign. She didn’t see it. She wasn’t intoxicated, and she wasn’t speeding. But she hadn’t driven this route before. And although she was watching the road, she was also thinking about her work; indeed, she had just realized how to solve a problem that had been bothering her for days. Her attention to her driving had dropped below where it ought to have been. (Clarke 2017: 238-39)

There is no need to imagine that Ann desires, values, or cares too little about driving safely, or desires, values, or cares too much about solving the work problem, or has an objectionable balance of these motivations and concerns, in order to understand how she could have failed in the way she did. She was, after all, attending to her driving—just not closely enough. Her attention was partly on something else—as ours commonly is when we drive—and given the unfamiliarity of her route, she missed seeing something that she didn’t expect to find. The failure is executive rather than motivational, a failure successfully to implement her preferences in action. Still, moral blame seems fitting. Ann has killed someone, and although she did not do so intentionally, she has no excuse.

Another kind of case involves failure to carry out a plan when one reverts to well-established routine.[[3]](#footnote-4) An example, **Hot Car**:

Consider a parent who leaves his baby in the parked car on a scorching summer morning. His intention was to deliver the baby to the child-care center, but he “just forgot” she was in the car. The baby had only begun child-care, so the drop-off was not part of an established routine. Moreover, it was the first time he had taken her with him on his way to campus. She slept quietly in the back seat the whole drive, while as usual he pondered his morning lecture. (Watson 2016: 399)

Each year numerous cases of this kind result in a child’s death from hyperthermia.[[4]](#footnote-5) Parents who make this fatal mistake are usually loving and, except for this incident, attentive to the child’s needs. There is neither a lack of motivation to care for the child nor excessive motivation to excel at work. What goes wrong is not motivational but executive: the motivation to deliver the child safely is not acted on as it should be.[[5]](#footnote-6) Still, blame seems to be merited, and it is moral blame. Parents who make this mistake feel terribly guilty. A plea for forgiveness is in order. A terrible wrong has been done, and neither ill will nor incapacity explains it.[[6]](#footnote-7)

In cases such as these, I contend, moral blame is merited, the agents are blameworthy for their failures, despite the fact that the failures do not result from objectionable quality of will. In the next three sections, I sketch a view of responsibility and blameworthiness that supports these judgments. In subsequent sections I respond to some objections.

1. Capacities and Competence

As rational agents, we have a capacity for reasons-responsive conduct.[[7]](#footnote-8) This capacity in turn consists of, or is subserved by, a number of others, some distinctly rational and others not. Our thought and action commonly manifest these several capacities. We form beliefs on the basis of evidence and by inference, usually taking ourselves to have good reason to believe what we believe. We notice and consider practical reasons, commonly appreciating them *as such*. We deliberate about what to do, anticipating outcomes of options under consideration, assessing and weighing reasons, sometimes judging options better, best, or good enough on various normative dimensions, and sometimes best overall. We organize our own agency through time and coordinate it with that of other agents. We exercise self-control, enabling us to resist temptation, sometimes employing skill and sometimes sheer will power. We make decisions and execute the intentions formed therein, acting when we do on the basis of our appreciation of reasons. For numerous kinds of action or activity, we possess the know-how required to carry out that conduct successfully.

The capacities manifested in doing the things mentioned so far are distinctly rational. But we can exercise them well only because we have certain other capacities that are not distinctly rational—their exercise is not an exercise of reason—and the exercise of some of which is not itself action. Perceptual capacities in various modalities are required if we are to recognize reasons (for example, that there is an obstacle ahead). Deliberation requires capacities for attention and concentration. Fruitful deliberation requires powers of imagination in conceiving of possible ends and means to those ends and in foreseeing consequences. It requires working memory; one can’t deliberate well if one can’t keep track of considerations that one had in mind just a moment ago. A capacity to think of things relevant to one’s deliberative task is required if one is to deliberate well. Strength, agility, and bodily coordination are needed to do many of the things that we are capable of doing.

These capacities wax and wane, as we gain or lose know-how, skill, strength, or mental acuity. They can be masked by fatigue or stress; they are dormant in sleep. But they are relatively stable features of human agents, varying in range and degree from one of us to another. One’s capacities of these kinds are part of what make one the person one is, and one’s understanding of one’s capacities of these kinds is central to one’s self-conception. We know that we have these capacities, we rely on their functioning well, and we know that we can so rely on them.

Joseph Raz uses the expression “domain of secure competence” to pick out a range of activities in which, given our capacities for thought and action, “we do not need to assess our chances of success before we take the action” (2011: 246). Within this range, we rightly take our proficiency for granted. Further, “our sense of who we are, which underpins our self-esteem, as well as our inclination to take or avoid risks and therefore our aspirations and ambitions, is tied up with our success in establishing a domain within which our powers of rational agency are securely reliable” (245). Raz draws the following connection to responsibility: “In holding ourselves competent within that domain, we hold ourselves responsible for actions that fall within it. To disavow responsibility for such action is to be false to who we are” (245).[[8]](#footnote-9)

Even within one’s domain of secure competence, one sometimes misjudges matters, fails to employ information that one possesses, fails to attend as one should, or forgets one’s plan. As we own our successes, so these failures are our own. As Raz puts it, conduct for which we bear responsibility “is conduct that is the result of the functioning, successful or failed, of our powers of rational agency, provided those powers were not suspended in a way affecting the action” (2011: 231).[[9]](#footnote-10) In a similar vein, Gary Watson says of the processes by which we manifest these capacities: “Their success is our success. By the same token, their failure is our failure” (2022: 117).

The proviso that Raz mentions is, of course, important. One’s capacities can be temporarily lost, blocked, or defeated by fatigue, sleep, pathology, or interference. A failure can be at least partly excused when due to such factors. But we own those that are not, when success was easy for us, even when our motivations are unobjectionable.

1. Blameworthiness

Raz’s topic is responsibility; it is not worthiness of blame, whether moral or otherwise. As he observes, “There is a whole variety of responses and attitudes mandated by responsibility, other than blaming or praising” (2011: 251).[[10]](#footnote-11) Angela Smith emphasizes a similar point. As she puts it (2015: 103), taking an agent to be responsible for something that she has done is finding her eligible for responses made appropriate by that conduct. It is a further question which responses are appropriate.

Although it is not Raz’s view,[[11]](#footnote-12) a position that might be taken about the failures of agency under consideration is that agents are responsible for them but never to blame, at least not non-derivatively. (The qualification allows that agents may be blameworthy for earlier conduct that foreseeably led to such a failure, and thus indirectly or derivatively to blame for the failure. Such indirect blameworthiness is not what is at issue here.) Here I consider a view of this sort.

Julie Tannenbaum (2015, 2018) argues for recognition of a deontic category that she calls “mere moral failure.” We have a case of such failure when an agent has a moral obligation, she does not satisfy that obligation, and her failure is not due to a blameworthy orientation of her will or any other blameworthy trait of hers, such as unreasonable overconfidence. In that case, Tannenbaum holds, although the agent fails to fulfill the obligation, she does not violate it, and her conduct is not morally wrong. Although she is responsible for the failure, she is not blameworthy for it. Examples that Tannenbaum takes to illustrate this status resemble several of those I have discussed.

I do not think that in the cases considered here it can be said that the agent does nothing morally wrong. Nor do I think that these agents are blameless.

Consider the matter from the perspective of victims in these cases. They have claims on others not to be harmed in the ways that they are. When these claims go unmet, they may properly take it that they have been wronged. And if they have been, it is the agents in these cases who have wronged them. Marie wrongs the pedestrian killed as a result of her misjudgment; Bob wrongs his neighbor by letting his son take the dog; Ann wrongs the passenger killed by her bad driving; the parent who leaves a child in a hot car wrongs the child.

In wronging others, these agents do wrong.[[12]](#footnote-13) Further, what agents do in these cases is not just objectively wrong. It is not just that what they do is, in fact, suboptimal, worse than an available alternative, whether they can know it or not, given the evidence available to them. What they do can be seen *by them* to be suboptimal. They have access to information, and they have capacities for judgment, choice, and action, such that alternatives that they know to be available can be seen by them to be better.

Tannenbaum maintains that “because the agents have all the regard and respect for others that they should have, and no blameworthy condition of theirs led to their action, their action is neither wrong nor blameworthy” (2018: 141). The claim presumes that conduct is wrong only if it is explained by a prior blameworthy fault in the agent. That is what I dispute. The moral wrongness of a failure does not, I contend, depend on whether it stems from a prior morally blameworthy state or condition.

Tannenbaum acknowledges that agents who fail in these ways may acquire duties of reparation, and that it can be appropriate for them to feel bad.[[13]](#footnote-14) However, she maintains, the appropriate emotional response is neither agent-regret nor guilt. Unlike cases in which the former is what is fitting, in these the agents had obligations that they failed to fulfill. And unlike cases in which it is the latter that is fitting, in these the agents did not do wrong. In fact, she holds, we have no name for this third kind of attitude that is fitting in the cases in question.

No doubt, our responses to agents in these cases should differ from responses to wrongdoers who act from ill will. As Tannenbaum says (2015: 66), the conduct of agents in these two kinds of case differs in kind, not just in degree. But blame—moral blame—can differ in a variety of ways. A blaming attitude can represent the agent has having acted from ill will, or not. Blame can be angry, or not. Still, some of the same kinds of reactive emotions that are fitting in cases of malicious wrongdoing can be fitting in cases of wrongdoing by goodwilled agents. A parent who forgets her child in a hot car can fittingly feel guilty; family members of someone killed by a driver who inadvertently ran a stop sign can be fittingly angry at the driver.

1. Moral Blameworthiness

All of the cases I have considered involve a failure of rationality of one kind or another. Marie has information in light of which it is irrational to judge it safe to keep the bridge open, yet that is the judgment she makes. Similarly, Bob’s decision to allow his son to take the dog is irrational in light of information that Bob possesses. Granted, that information does not come to mind when Bob deliberates about the matter. But that itself is a failure of rationality, a failure to employ in deliberation pertinent knowledge that he possesses. Ann’s failure to remain vigilant in looking for traffic signs is irrational given the value that she places on driving safely; her failure to attend is a failure to act as she values acting. The parent who unwittingly leaves a child in a parked car at work irrationally fails to execute an intention that he has, and he does something that he very much prefers not to do.

More broadly, the failures in these cases occur in what may fairly be called the will. They all concern personal-level attitudes that can figure in deliberation and decision. There is a fault of will in these cases, though it is not a deficiency in desires, values, or cares. The agents’ concerns are as they should be; the agents fail to manifest their concerns in action, and nothing excuses their failures. Failures of agency in which there is worthiness of moral blame, I suggest, include those in which there is this kind of fault of will, as well as those in with there is objectionable quality of will.

Capacitarianism, Narrower and Broader

The view of responsibility and blameworthiness that I have suggested is capacitarian. The range of our responsibility, and of things for which we may be blameworthy, depends on our capacities for thought and action. Capacities that constitute and subserve our general capacity for reasons-responsive conduct include some that are specifically rational and others that are not. I have suggested that we can be blameworthy even when our conduct is due to failure to manifest a capacity of the latter kind, when the resulting conduct is nevertheless due to some fault of will. Some of these failures are failures to manifest capacities that are not distinctly rational, such as failure to remember one’s plan.[[14]](#footnote-15)

A narrower capacitarian view limits blameworthiness to cases in which failure to do the right thing stems from failure to manifest a specifically rational capacity. One recognizes what are in fact reasons to desire to do a certain thing, but one does not respond with the appropriate desire. One reasons poorly in forming a belief. Or one rationally judges a certain alternative best but fails to intend in accord with one’s judgment. Such a narrower capacitarian view might find Marie blameworthy for the death she causes but not find blameworthy the parents who forget their children in their cars. (I am not sure what it would say about Bob and Ann.)

One consideration that might push a capacitarian toward this narrower view is the worry that some failures—such as the father’s failure to remember to drop his child at day care—are mere “glitches,” breakdowns in the functioning of processes that are not under one’s rational control and are therefore not one’s fault. Like the well-intentioned agent who drops a valuable vase, one who forgets has failed but is not to blame.[[15]](#footnote-16)

The failure to transition from a standing plan to a present-directed intention at the appropriate moment is a failure of rationality, even if memory is not itself a rational capacity. And the worry about glitches can be pressed against a narrower capacitarianism as well. Even Marie’s misjudgment might be characterized as due to a glitch in a type of process not under her direct volitional control.[[16]](#footnote-17) Though her capacity to judge the relevant matters usually serves her well, it did not this time, and the failure is (in Tannenbaum’s terms) “simple,” in the sense that it stems from no moral deficiency in her.[[17]](#footnote-18)

To attribute failure to a glitch is already to suggest that it is not the agent’s fault. One might add that it is just bad luck that these agents’ faculties failed them on these occasions. But we do not reject credit for the successes when luck is on our side. It is unclear what might justify the rejection of blame when it is not.[[18]](#footnote-19)

Some failures to manifest our agential capacities can be attributed to factors that mask or temporarily diminish them. Stress can mask one’s capacity to focus on relevant matters. Narcotics mask or diminish one’s mental capacities. Such factors commonly provide at least partial excuses when they lead to failures of agency. In contrast, although “I forgot,” “I wasn’t paying attention,” “I didn’t think of that,” and “I misjudged matters” might rule out malicious or deliberate wrongdoing, they are not excuses.

The view advanced here is not the broadest kind of capacitarianism. It would not find an attentive and well-motivated agent morally blameworthy for dropping a vase.[[19]](#footnote-20) The failures for which one can be to blame, on this view, do not include such pure performance errors. On a spectrum of capacitarian views, this one lies between extremes. It is a Goldilocks view.

A Different Kind of Blameworthiness?

I’ve argued that agents in the cases considered here are blameworthy for their failures. Some might agree but insist that it is not an ordinary kind of moral blameworthiness that the agents bear. Elinor Mason, for example, identifies what she calls “extended blameworthiness.” In some cases of inadvertent wrongdoing, she holds, agents do not bear ordinary blameworthiness, but they should *take* responsibility for their conduct and its consequences. Taking responsibility, in the intended sense, is not a matter of accepting something that holds independently of one’s acceptance. Rather, “in taking responsibility…[agents] become properly blameworthy in the extended way” (2019: 180).[[20]](#footnote-21)

On Mason’s view, in the cases in question, the agent bears blameworthiness for her misdeed only if, subsequent to it, she takes responsibility for it. The agent “can avoid blameworthiness by refusing to take it on” (2019: 201). I am skeptical that there is a kind of moral blameworthiness that has this feature. In any event, blameworthiness borne by the agents in the cases I have discussed does not. Whether or not these agents take responsibility, they are to blame.

Gunnar Björnsson observes that there is “skill blame” that can be appropriate when moral blame is not.

Noticing and locating objects based on visual information, remembering things, or solving mathematical problems are activities that can be performed more or less skillfully, with greater or lesser successes and failures. However, while you might kick yourself for failing to solve a mathematical puzzle, characteristically *moral* modes of blame involving guilt and indignation seem inappropriate. (2017b: 143)

Failures merit moral blame, he maintains, only when they are due to moral shortcomings of the agents, in particular, to substandard quality of will. An agent’s quality of will is “a matter of how well she *cares* about what is morally important” (149). Wrongdoing that is not explained by a fault in the agent’s cares is not something for which she merits moral blame.

I do not think that this jibes with how we judge these matters. Agents in the cases in question have not made mistakes on math quizzes or GRE exams. They have killed or caused serious injury to other persons. They did so because they failed to exercise capacities for rational action that they possessed to do things that were well within their domains of competence. There is nothing in themselves or their circumstances that excuses their wrongdoing.

Error Theories

I have argued that agents can be morally blameworthy for failures that do not result from objectionable quality of will, where the latter is said to concern one’s desires, values, or cares. I contend that everyday blaming responses and judgments of blameworthiness agree with this view. Some authors accept the latter claim but offer error theories of these judgments.

Matthew Talbert responds in this way in defense of a quality-of-will requirement for blameworthiness. He considers cases with the same structure as those of good-willed parents who unwittingly leave their children in hot cars. Talbert tentatively offers the following explanation of the supposed error of finding these agents blameworthy:

Given the importance (for social beings such as ourselves) of being aware of how other members of our community are oriented toward us, perhaps we are primed to recognize instances in which we have been treated with inappropriate regard. Indeed, we may be so vigilant in this domain that we tend to make erroneous attributions of disrespect. Cases of unwitting omission [like some considered here] would be well suited for triggering such misattributions since they involve harm (which often *is* the result of disregard) and the genesis of the harm is often unclear. Perhaps, because of our interest in uncovering signs of disrespect, we tend to explain harm in these cases in ways that cast the supposed cause of the harm in a morally worse light, at least when evidence to the contrary is absent. (2017: 31)

Although it is no part of the presentation of the cases that the agents have objectionable desires, values, or cares, given the harm that results, Talbert suggests, we infer that they do. Our attributions of blameworthiness in these cases, then, are based on mistaken attributions of bad attitudes. They may thus be dismissed as erroneous.

Murray et al. (2019) report results that place Talbert’s conjecture in doubt. They conducted three studies collecting responses to vignettes about an agent forgetting to do something he had promised to do. Study 1 asked for agreement or disagreement with statements that the omission was the agent’s fault, that it was fair to blame him, and that he should feel guilty; studies 2 and 3 asked how much fault the agent had, how much he should be blamed, how much guilt he should feel, and how much moral responsibility he had for the omission. One variable manipulated across the vignettes was the agent’s level of care—how much he cared about doing what he had promised. (Level of stress was also manipulated.) Even for high-care conditions, mean ratings were generally well above the mid-point of the 7-point scale for fault, fairness of blame, appropriateness of guilt, and responsibility. Effect sizes of care in studies 1 and 2 were small; in study 3 there was no significant effect of care on judgments of fault and blame, and only minimal significant effects on judgments of guilt and moral responsibility. Participants correctly reported the level of care indicated in the vignettes, showing no confusion on this matter. Discussing the results, the authors suggest that “care plays a minimal role in moral assessment” (1196). Apparently, then, people commonly blame agents despite recognizing that their wrongdoing does not stem from bad attitudes.[[21]](#footnote-22)

Neil Levy considers **Bad Dog** and **Stop Sign**, as well as a case structurally like **Hot Car**,[[22]](#footnote-23) offering a rather different error theory of the judgments I have offered on these cases. He maintains that our disposition to blame is an adaptation, a trait favored by natural selection because of the enhancement to fitness that it (or its precursors) conferred. Survival of human beings, Levy observes, depends on cooperative interaction. But under a wide range of conditions, defection is a dominant strategy: it is advantageous to refuse to cooperate, regardless of whether others do. Punishment of defectors can change expected outcomes and thereby secure cooperation, and thus its benefits. “Blame judgments,” Levy suggests, “are designed to signal appropriate targets of punishment” (2017: 258) to secure social cooperation. If this is correct, he continues, “we ought to be disposed to blame those individuals who are likely to be bad bets for cooperative enterprises: those who intentionally defect from these enterprises” (258). Individuals who accidentally fail to cooperate “should not be blamed, because their doing so is not evidence of a disposition to defect in the future” (258).

Given the evolutionary history of our disposition to blame, Levy proposes, “we are committed to an implicit theory of moral responsibility, according to which agents are (directly) blameworthy only for those actions which they believe or desire will cause harm” (260).[[23]](#footnote-24) We are thus disposed in assigning blame to track agents’ beliefs and desires. Since these can be hard to discern—and defectors may try to deceive us about them—we commonly rely on conduct and its consequences for evidence. But such evidence can be misleading, as when someone unwittingly fails to conform to a cooperative arrangement. In such cases, Levy maintains, evidence provided by conduct and its consequences can trigger our disposition to blame.

In these kinds of circumstances, when we have good evidence that the agent’s mental states were not indicative of a bad quality of will, attributing blame to the agent is an error: it conflicts with our implicit theory. But we will find ourselves feeling that we ought to blame. The folk can be expected to blame in these circumstances, even though they are implicitly committed to an account of what justifies blame that entails that they ought not to blame. (2017: 261)

People may persist in such errors even when they explicitly accept that the wrongdoers lack objectionable attitudes, thus maintaining judgments that are mistaken “*by their own lights*” (261-62). (With this last claim Levy’s error theory departs from Talbert’s.)

These conjectures raise a number of interesting questions, most of which I must leave aside here. I limit my response to two points. The first concerns the assimilation of mechanisms of defector detection, punishment, and blame. The second concerns the normative import of adaptive dispositions.

Levy offers a just-so story of the natural selection of what might be called a defector detection mechanism that, he says, disposes us to punish.[[24]](#footnote-25) He suggests that our disposition to blame is explained by the same selection process. But defection can be deterred by refusal to cooperate with defectors, and a refusal to cooperate need not be an imposition of punishment. Punishment has an aim, one of imposing a cost in return for alleged misconduct, that such refusal can lack, for the refusal might aim only at protecting oneself from further loss. A mechanism of detecting defectors, then, is distinct from a disposition to punish.

Further, the latter is not the same as a disposition to blame. We blame without punishing, and we punish without blaming. We commonly blame the dead, though they are not available to be punished. We recognize quite different norms for the two. Offenses that merit blame can be too minor to merit punishment. And while nearly anyone may blame (perhaps the hypocritical are excepted), only those in special positions may punish. Even supposing, then, that a mechanism that produces punishment and one that produces blame are adaptations, it is questionable whether we have here a single mechanism with a single adaptive function (and a single set of norms).

A second (and I think more important) response concerns the normative lesson that Levy draws. Suppose, as he maintains, that we possess, as an adaptation, a disposition to blame that embodies an implicit theory. What, if anything, is the normative authority of that theory? Perhaps: if our blaming is to serve the purpose for which the disposition to blame (or its precursors) evolved, we must blame those who are such-and-such and not those who are so-and-so. But conditioned as it is on serving this evolutionary function, the norm has no moral import. We do not generally do moral wrong by failing to employ our capacities for purposes for which they evolved, nor for employing them for purposes quite different from those for which they evolved. (Celibacy is not immoral, and there is nothing morally amiss with doing a handstand.) For moral guidance on blaming, we shall have to look elsewhere.

Levy apparently agrees, for he rejects the theory that he finds implicit in our disposition to blame. He reports: “I am a moral responsibility skeptic: I do not believe that we ought *ever* to blame” (261).

The fact (if it is a fact) that our disposition to blame can be explained by its (or its precursors’) contribution to fitness does not settle the relevant normative matters. There is no such short cut from biology to morality.

Conclusion

We often fail at things that are relatively easy for us to do. Sometimes such failures result from our having objectionable attitudes, but often they do not. We commonly blame agents even for failures of the latter kind, and on reflection find these agents worthy of blame. I sketched a capacitarian view of responsibility on which such judgments can be correct. Failures to do what is easy for us to do can be failures to manifest capacities that constitute and subserve our general capacity for reasons-responsive agency. When we lack excuse, we can be to blame even for such failures that do not result from ill will, when in failing we wrong someone. Our conduct can then be morally wrong. These cases, I suggested, are limited to those in which our failure is in one way or another a failure of rationality and a fault of will.

I considered and rejected two error theories of blame, and of judgments of blameworthiness, in the cases in question. It does not appear that these responses stem from misattribution of bad attitudes to agents. And a conjectured evolutionary explanation of our disposition to punish fails to show that blame and judgments of blameworthiness are mistaken in these cases.

Often when we are to blame, our conduct is due to malice or indifference. But this is not so in every case. Even good-willed agents can be blameworthy for their failures.[[25]](#footnote-26)

**References**

Amaya, Santiago. 2013. “Slips.” *Noûs* 47: 559-76.

Amaya, Santiago and John M. Doris. 2015. “No Excuses: Performance Mistakes in Morality.” In

*Handbook of Neuroethics*, ed. Jens Clausen and Neil Levy, 253-72. Dordrecht: Springer.

Arpaly, Nomy and Timothy Schroeder. 2014. *In Praise of Desire*. New York: Oxford University

Press.

Ayers, Alisabeth. 2021. “Blaming for Unreasonableness: Accountability Without Ill

Will. *Journal of Ethics and Social Philosophy* 19: 56-79.

Björnsson, Gunnar. 2017a. “Explaining (Away) the Epistemic Condition on Moral

Responsibility.” In *Responsibility: The Epistemic Condition*, ed. Philip Robichaud and

Jan Willem Wieland, 146-62. Oxford: Oxford University Press.

------. 2017b. “Explaining Away Epistemic Skepticism about Culpability.” *Oxford Studies in*

*Agency and Responsibility* 4: 141-64.

Clarke, Randolph. 2015. “Negligent Action and Unwitting Omission.” In *Surrounding Free Will:*

*Philosophy, Psychology, Neuroscience*, ed. Alfred R. Mele, 298-317. New York: Oxford

University Press.

------. 2017. “Ignorance, Revision, and Commonsense.” In *Responsibility: The Epistemic*

*Condition*, ed. Philip Robichaud and Jan Willem Wieland, 233-51. Oxford: Oxford

University Press.

------. 2020. “It’s Up to You.” *The Monist* 103: 328-41.

Cosmides, Leda and John Tooby. 1992. “Cognitive Adaptations for Social Exchange.” In *The*

*Adapted Mind: Evolutionary Psychology and the Generation of Culture*, ed. Jerome H.

Barkow, Leda Cosmides, and John Tooby, 163-228. New York: Oxford University Press.

Fischer, John Martin, and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral*

*Responsibility*. Cambridge: Cambridge University Press.

FitzPatrick, William J. 2008. “Moral Responsibility and Normative Ignorance: Answering a New

Skeptical Challenge.” *Ethics* 118: 589-613.

Frick, Johann. ms. “Dilemmas, Luck, and the Two Faces of Morality.”

https://www.academia.edu/41957601/Dilemmas\_Luck\_and\_the\_Two\_Faces\_of\_Morality

Levy, Neil. 2017. “Methodological Conservatism and the Epistemic Condition.” In

*Responsibility: The Epistemic Condition*, ed. Philip Robichaud and Jan Willem Wieland,

252-65. Oxford: Oxford University Press.

Lloyd, Elisabeth A. 1999. “Evolutionary Psychology: The Burdens of Proof.” *Biology and*

*Philosophy* 14: 211-33.

Mason, Elinor. 2019. *Ways to Be Blameworthy: Rightness, Wrongness, and Responsibility*.

Oxford: Oxford University Press.

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

Murray, Samuel, Elise D. Murray, Gregory Stewart, Walter Sinnott-Armstrong, and Felipe De

Brigard. 2019. “Responsibility for Forgetting.” *Philosophical Studies* 176: 1177-1201.

Nelkin, Dana Kay, and Samuel C. Rickless. 2017. “Moral Responsibility for Unwitting

Omissions: A New Tracing View.” In *The Ethics and Law of Omissions*, ed. Dana Kay

Nelkin and Samuel C. Rickless, 106-30. New York: Oxford University Press.

Railton, Peter. 2009. “Practical Competence and Fluent Agency.” In *Reasons for Action*, ed.

David Sobel and Steven Wall, 81-115. Cambridge: Cambridge University Press.

Raz, Joseph. 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.

Reis-Dennis, Samuel. 2018. “Responsibility and the Shallow Self.” *Philosophical Studies* 175:

483-501.

Rosen, Gideon. 2015. “The Alethic Conception of Moral Responsibility.” In *The Nature of*

*Moral Responsibility: New Essays*, ed. Randolph Clarke, Michael McKenna, and Angela

M. Smith, 65-87. New York: Oxford University Press.

------. 2022. “The Problem of Pure Negligence.” In *Agency, Negligence and Responsibility*, ed.

Veronica Rodriguez-Blanco and George Pavlakos, 15-36. Cambridge: Cambridge

University Press.

Smith, Angela M. 2015. “Responsibility as Answerability.” *Inquiry* 58: 99-126.

Strabbing, Jada Twedt. 2019. “Accountability and the Thoughts in Reactive Attitudes.”

*Philosophical Studies* 176: 3121-40.

Strawson, Peter. 1962. “Freedom and Resentment.” *Proceedings of the British Academy* 48: 187-

211.

Talbert, Matthew. 2017. “Omission and Attribution Error.” In *The Ethics and Law of Omissions*,

ed. Dana Kay Nelkin and Samuel C. Rickless, 17-35. New York: Oxford University

Press.

Tannenbaum, Julie. 2015. “Mere Moral Failure.” *Canadian Journal of Philosophy* 45: 58-84.

-----. 2018. “Moral Responsibility without Wrongdoing or Blame.” *Oxford Studies in Normative*

*Ethics*, vol. 8, ed. Mark C. Timmons, 124-48. Oxford: Oxford University Press.

Watson, Gary. 2016. “Raz on Responsibility.” *Criminal Law and Philosophy* 10: 395-409.

------. 2022. “The Possibility of Pure Negligence.” In *Agency, Negligence and Responsibility*, ed.

Veronica Rodriguez-Blanco and George Pavlakos, 99-117. Cambridge: Cambridge

University Press.

Weingarten, Gene. 2009. “Fatal Distraction: Forgetting a Child in the Backseat of a Car Is a

Horrifying Mistake. Is It a Crime? *Washington Post Magazine*. March 8.

1. Proponents include Arpaly and Schroeder (2014: ch. 7), Björnsson (2017a, 2017b), McKenna (2012: ch. 3), Rosen (2015), Strabbing (2019), and Talbert (2017). Proponents often cite Strawson (1962) as inspiration for a quality-of-will view. Rosen declares that almost all writers who view moral responsibility, as Strawson did, as liability to the reactive emotions accept that “*X* is blameworthy for *A* only if in doing *A* *X* showed an objectionable pattern of concern” (2022: 16). I find it questionable whether Strawson held such a view; but, of course, whether he did is not to the point. [↑](#footnote-ref-2)
2. FitzPatrick (2008) argues that we can be responsible for ignorance that manifests such vices. If the view I develop here is correct, then we can blameworthy for failures when neither epistemic vices of this sort nor objectionable quality of will figures in the etiology. [↑](#footnote-ref-3)
3. “When people fail to monitor their performance adequately, their behavior will automatically tend to follow well-rehearsed routines consistent with their immediate surroundings and their general know-how. As a consequence, what they wind up doing is not necessarily in line with what they intended to do, or the reasons that first motivated them to act” (Amaya and Doris 2015: 265). [↑](#footnote-ref-4)
4. For a perceptive report on such cases, see Weingarten (2009). [↑](#footnote-ref-5)
5. Amaya (2013) argues that mistakes of this kind fail to reveal the agents’ preferences. The parents at no point prefer driving straight to their ultimate destination to first dropping off the child. Rather, they fail to act as they prefer. [↑](#footnote-ref-6)
6. While rejecting the appropriateness of criminal punishment in such cases, Watson observes: “there is a strong case for forgiveness by all those concerned. Nevertheless, there is something to forgive: the negligent parents who are plagued by guilt (and shame) are not necessarily making a normative mistake, it seems to me” (2016: 400). There is also a case for mercy; the parents, suffering the loss of a child, are often racked with guilt. [↑](#footnote-ref-7)
7. This paragraph and the next draw from Clarke 2020. Compare the inventory of capacities said to constitute what Railton (2009) calls “practical competence.” [↑](#footnote-ref-8)
8. I have omitted the subscript that Raz places on ‘responsible’. He uses it to indicate responsibility for an action, omission, or consequence that presupposes that one is a responsible person, that is, one possessing the capacities of rational agency. Such responsibility is also distinguished from that which amounts to having a duty to do a certain thing (“It is your responsibility to water the plants each day”). See his 2011: 227-28. [↑](#footnote-ref-9)
9. It is direct or non-derivative responsibility that is Raz’s concern here, responsibility for conduct that does not derive from one’s responsibility for prior conduct. [↑](#footnote-ref-10)
10. Again, I have omitted Raz’s subscript (see note 8). [↑](#footnote-ref-11)
11. About blame, Raz says, “Agents are to blame for conduct for which they are responsible unless they are excused” (251; again I omit his subscript). Apparently, he is considering here only cases of wrongdoing. Of course, we can be *praiseworthy* for conduct for which we are responsible, and being praiseworthy is not a matter of having an excuse. [↑](#footnote-ref-12)
12. It is sometimes argued (for example, by Frick ms) that an agent might wrong someone but not act wrongly, if she fails to fulfill a special obligation she has to that person while doing what is all things considered morally permitted or required. But the cases here are not of this kind. [↑](#footnote-ref-13)
13. Tannenbaum discusses whether one or another emotional response is “justified.” She offers different glosses on this normative status. At one point, she says, “if a feeling is justified, then the content of the thought that constitutes or is associated with the emotion is correct (as well as the emotion’s duration, intensity and so on)” (2015: 77 n. 6). Correctness of the thought would (presumably) be its truth. However, elsewhere she says, “an emotion is justified only if the belief that is conceptually associated with the emotion is reasonably believed” (2018: 125). Since what is true and what is reasonably believed can differ, the two construals differ. And certainly the second does not correlate with blameworthiness; one can be reasonably believed to have committed an offense of which one is innocent. I take fittingness to be the normative status that blame must have when one is blameworthy. The first of the construals of justification offered by Tannenbaum might be taken to be view of it as fittingness. However, I don’t rely here on any particular view of what fittingness is. [↑](#footnote-ref-14)
14. The account advanced here is closely related to the reasons-responsiveness view set out in Fischer and Ravizza (1998). One difference: whereas my account is stated in terms of *agents’* capacities, Fischer and Ravizza’s view is stated in terms of the responsiveness of action-producing *mechanisms*. Fischer and Ravizza do not address cases of failure like those discussed here, and I am not sure what view they would take of such cases. [↑](#footnote-ref-15)
15. Ayers (2021: 75) raises this objection to broader capacitarian views in advancing her narrower “Rational Capacitarianism.” [↑](#footnote-ref-16)
16. Rosen (2022), in arguing for a quality-of-will view, rejects blameworthiness in many cases of negligence on the ground that the failures to advert are due to glitches. [↑](#footnote-ref-17)
17. “A lapse is simple just in case there is nothing of moral relevance to say about what led to the lapse” (Tannenbaum 2018: 133). [↑](#footnote-ref-18)
18. It might be said that the successful operation of our faculties within our domain of competence is overwhelmingly likely, and hence not lucky. But credit can be due even for lucky successes. [↑](#footnote-ref-19)
19. Another case in which it would not find the agent morally blameworthy: Blair Walsh’s missed field goal attempt (as discussed in Reis-Dennis 2018: 495). [↑](#footnote-ref-20)
20. The cases in which Mason says agents should take responsibility for inadvertent wrongdoing mostly involve close personal relationships. Even in these, she holds, taking responsibility is not a duty; it is morally optional. [↑](#footnote-ref-21)
21. A further point: as Nelkin and Rickless observe (2017: 110), parents who forget their children in cars with fatal effect typically blame themselves severely, but presumably without mistakenly attributing bad attitudes to themselves. [↑](#footnote-ref-22)
22. He also considers cases from Clarke 2015 and 2017 that I have not discussed here. [↑](#footnote-ref-23)
23. Here Levy states the implicit theory as holding that we are directly to blame *only for* actions we believe or desire will cause harm. Earlier in the same paper (2017: 253) he states it as holding that we are directly blameworthy *for* actions we perform intentionally. Setting aside the difference in the characterization of the actions, the two statements together apparently characterize a theory on which we are directly to blame *for* *and only for* actions of a certain kind. I read him as attributing such a theory to us. [↑](#footnote-ref-24)
24. This defector detection mechanism is, apparently, the cheater detection mechanism hypothesized by Cosmides and Tooby (1992). For critical assessment of their proposal, see Lloyd (1999). [↑](#footnote-ref-25)
25. I am grateful to the editors of this volume for comments on earlier drafts of this chapter. [↑](#footnote-ref-26)