Psychopaths, Incorrigible Racists, and the Faces of Responsibility*

Dana Kay Nelkin

Abstract: Psychopaths pose a puzzle. The pleasure they take in the pain of others suggests that they are the paradigms of blameworthiness, while their psychological incapacities provide them with paradigm excuses on plausible accounts of moral responsibility. I begin by assessing two influential responses: one that claims that psychopaths are morally blameworthy in one sense and not in another, and one that takes the two senses of blameworthiness to be inseparable. I offer a new argument that psychopaths, as understood in the debate, are blameworthy in neither sense, while showing how the two senses of blameworthiness nevertheless come apart.

Do I care about other people? That’s a tough one. But, yeah, I guess I really do…but I don’t let my feelings get in the way…I mean, I’m as warm and caring as the next guy, but let’s face it, everyone’s trying to screw you…Do I feel bad if I have to hurt someone? Yeah, sometimes. But mostly it’s like…uh…[laughs]…how did you feel the last time you squashed a bug? (Quote from a diagnosed psychopath in Hare, *Without Conscience*)

Psychopaths pose a puzzle. On the one hand, they seem to be the paradigms of evil and blameworthiness. On the other hand, as we learn more about the emotional and cognitive deficits of psychopaths, we realize that they lack certain capacities that we ordinarily associate with responsible agency. How can agents who are constitutionally incapable of seeing other people’s pain as a reason to refrain from hurting them be blameworthy for their actions? In this way, they seem more like wild animals who simply don’t have the capacities that underlie responsible agency. Hence, the puzzle. When we think about psychopaths’ actions, and the
pleasure they sometimes take in inflicting harm, it seems we are correct in blaming them; when we think about their deficits, it seems we are incorrect to do so.

At first glance, it appears that we have to give up something that strikes us as compelling. But another approach that appears to leave a way out is available: it is to see that our reactions are in tension only if we equivocate between two different notions of responsibility, one that concerns the attribution of virtues and vices, another that concerns the appropriateness of holding responsible, making demands, and responding with reactive attitudes such as resentment and indignation. Once we correctly distinguish between the different notions—one with which we rightly respond to the cruelty apparently displayed by psychopaths and one with which we wrongly hold them accountable for their behavior, our puzzle dissolves. (Following Watson, who distinguishes between two “faces” of responsibility, let us call the first notion the “attributability” face and the second the “accountability” face.) And at the same time, psychopaths would seem to be the perfect illustration of the distinction, thereby offering additional support to the idea that there is a real distinction to be made. I will call this solution to the puzzle, which includes Watson’s particular way of distinguishing between two faces of responsibility, and its application to the psychopath, the “Separable Middle Course View,” or “Middle Course View” for short.

Despite the elegance of this solution, a quite different approach has recently received a spirited defense. On this alternative approach, psychopaths are not only responsible in the sense of being properly attributed with vices such as cruelty, but they are also properly held accountable, made subject to demands, and responded to with reactive attitudes. Further, and crucially, what psychopaths illustrate is not that there are two distinct notions of moral
evaluation that come apart, but--to the contrary--that the first is sufficient for the second. To capture both of these commitments, I will call this view the “Unified Blameworthiness View”.

I believe that each offers great insight, but neither is correct about all key claims. In the end, I argue that there are important distinctions to be made between different notions associated with responsibility—on this point agreeing with the Middle Course View—but also that psychopaths are not the best illustration—and on this point disagreeing with it. In so doing, I offer a new argument against an assumption shared by both the Middle Course View and the Unified Blameworthiness View, namely, that moral vices such as cruelty are properly attributed to psychopaths. At this point, the more general question about how the two kinds of moral appraisal are related remains, and I go on to show how disambiguation of the notion of “demands” can help answer this fundamental question. Ultimately, I offer a view that departs from the Middle Course View not only in its treatment of the psychopath, but also in its particular understanding of the different notions associated with moral responsibility. For this reason, I call the view I defend the “Separable No Blameworthiness View”.³

Of course, every party to the debate is making certain assumptions about the nature of psychopaths, assumptions that may be undermined by further research into those we have up to now characterized as psychopaths. In section 1, I spell out, and then take on board until the final section, a common conception of the deficits of psychopaths. In sections 2 and 3, respectively, I describe the Middle Course View and the Unified Blameworthiness View in more detail. In section 4, I then offer a new argument against the common assumption, made by both views, that the psychopath is blameworthy in the attributability face and properly described as in possession of moral vices such as cruelty. Abstracting from the psychopath in section 5, I defend one aspect of the Middle Course View, namely, that attributability is insufficient for accountability. In
sections 6 and 7, I consider powerful objections to this thesis from the Unified Blameworthiness View. In section 8, I bring together a number of theses defended in previous sections to show how the view I defend departs from the Middle Course View not only in its application to the psychopath, but also in its conception of attributability and accountability. I then come full circle, returning to the question of whether the original view of the psychopath’s deficits is correct, how we might proceed in finding out, and what is at stake in so doing.

1. Some assumptions about the psychopath

Interestingly, psychopathy has never been a category in the Diagnostic and Statistical Manual, although it was considered for inclusion in the recently released DSM-V. Psychopathy is mentioned there in the discussion of the diagnostic features of Antisocial Personality Disorder, where it is claimed that “[T]his pattern has also been referred to as psychopathy, sociopathy, or dissocial personality disorder”. However, given the disjunctive set of diagnostic criteria for Antisocial Personality Disorder, one could easily meet the conditions for this diagnosis without fulfilling what have to this point been taken to be defining traits of psychopathy in previous diagnostic tests. Currently, the main test is the Hare Checklist, a two-factor model that takes into account “interpersonal/affective” items and behavioral dispositions, or “impulsive/antisocial lifestyle items.” Notably, it includes “lack of remorse or guilt” and “callous/lack of empathy,” emotional deficits that have been thought to be at the root of the disorder. In addition to their performance on the Hare Checklist, there has been much discussion in the literature of psychopaths’ weak performance on a well-known test for moral cognition, Turiel’s moral/conventional test, that shows whether subjects distinguish moral transgressions from
conventional ones (like using the wrong fork for a salad course). The test is commonly thought to measure moral knowledge.

It is notable that the diagnostic tools themselves are not presented as sets of necessary and sufficient conditions for psychopathy. The Hare Checklist, for example, yields scalar results, and two subjects might receive the same total score by satisfying quite different conditions on the list. For this reason, and because the construction of tools themselves is in progress, it makes sense that philosophers and legal theorists not rely exclusively on a single measure as capturing perfectly the features of psychopathy.

Instead, philosophers tend to stipulate certain features that seem to be central to the condition and that have great potential interest for moral and legal theory independently of when or even whether they are ever instantiated. For example, Matthew Talbert writes,

I will not take up the dispute about how to divide the terrain of personality disorders. Thus, my usage will be stipulative and my characterization of the disorder will be very brief. If a reader finds that my account of psychopathy does not fit her understanding of the disorder, she may wish to substitute ‘morally blind agent’ for ‘psychopath’ in what follows…I wish to focus here on moral blindness as a cognitive deficiency…Psychopaths, as I shall conceive them, are morally blind agents, unable to grasp, for instance, the difference between moral requirements and requirements generated by social conventions. This failure of moral understanding explains why psychopaths are unmotivated by moral considerations that are salient to other people.

And Gary Watson writes,
…psychopathy (as I read the evidence) precisely involves an incapacity to recognize the interests of others as making valid claims on them. Consequently, they are disabled from standing in the reciprocal relations or (to use another idiom) from engaging in the mutual recognition that lies at the core of moral life. In John Rawls’ sense, they lack the features of moral personality: a sense of justice and a conception of the good.⁸

Both Talbert and Watson offer very different solutions to the puzzle of psychopaths with which we started, but their initial description of psychopaths is similar. Both focus on the central feature of (lack of) receptivity to moral reasons and to the lack of understanding that someone else’s interests provide non-instrumental reasons for acting. In the next sections, I will take this common conception of psychopathy on board. In the final section, I will return to the question of how we can tell whether the conditions in question are actually instantiated.

2. The Middle Course View: Two Faces of Responsibility Distinguished and Applied

In a set of influential papers, Watson describes two “faces” of responsibility, and offers an elegant application of this distinction to solve our puzzle about psychopaths.

One might blame someone in the sense of attributing a moral fault to an agent. One might call a piece of behavior (e.g., a colleague’s failure to return a book) “shoddy,” for example. So far, this is only to blame in the “attributability” sense. But this is not yet to blame in a sense that presupposes that the agent “deserves adverse treatment or ‘negative attitudes’ in response to their faulty conduct” (Watson, “Two Faces,” 266). To do the latter is to blame in the “accountability” sense.
To elaborate, one is responsible in the attributability sense if one is appraisable insofar as one’s actions reflect one’s having adopted an end or insofar as one stands for something. As Watson points out, since “many of these appraisals concern the agent’s excellences and faults—or virtues and vices—as manifested in thought and action,” judgments like those that attribute shoddiness are made from the “aretaic perspective” (ibid., 265-66). Thus, one is responsible for an action in the attributability sense when it is correct to make this sort of aretaic judgment. In contrast, our practices of holding one another accountable “involve the imposition of demands on people” (ibid., 273). In turn, imposing demands is a matter of laying it down that “unless the agent so behaves she will be liable to certain adverse or unwelcome treatment” or sanctions. Sanctions are in turn connected with the reactive attitudes like indignation, resentment and disapprobation in the following way: either they are themselves disagreeable when experienced by their targets, or they involve dispositions to treat others in generally unwelcome ways. Because the prospect of adverse treatment arises in this way, questions of fairness arise in connection with accountability, and it is for this reason that many have thought being able to do otherwise is essential for accountability. In order to be blameworthy in the accountability sense, the demands in question must be fair or just.

Now, for Watson, demands play a central role in accountability. But the content of the demands is not yet clear. If a demand is thought to be expressed as part of a blaming response that can only be expressed after the action in question is performed, then it might appear that it is a demand relating to future behavior, such as a demand not to act the same way again. (It wouldn’t make sense to now demand that you act differently two days ago.) Perhaps such future-directed demands are part of our blaming practices. But, in my view, insofar as one is truly accountable for a past action, it is because one was at the time of the action subject to a
demand--a demand to act differently at that time, whether that demand was actually expressed or not. I will return to the question of the particular contents of the relevant demands in section 5. For now, importantly, the crucial idea that I take to be plausible and that I will defend is that one is accountable when one is subject to a demand at the time of action.

With the distinction in hand, we can see how it can be applied to the psychopath. As Watson writes,

We rightly predicate viciousness of the attitudes and conduct of psychopaths (“dishonest,” “abusive,” “manipulative,” “hostile,” “mean”), and we naturally respond to them accordingly. This is the attributability face. And yet they lack the capacity for moral reciprocity or mutual recognition that is necessary for intelligibly holding someone accountable to basic moral demands and expectations. This capacity, in my view, delineates the boundaries of moral accountability, and thereby the other face of responsibility. (Watson, “Trouble,” 308)

Thus, psychopaths are responsible in the attributability face, but they fall short of accountability. The reason for their failure to be accountable is that, as Watson puts it, it is not conceptually apt to make a demand of a creature that is incapable of recognizing one’s standing to make demands. Because accountability is a matter of the aptness of making demands, psychopaths simply cannot qualify.

At the same time, it seems that they do satisfy the conditions for attributability. As Watson puts it,
psychopaths are not just dangerous but cruel. They frequently enjoy forcing others into submission, for example…That psychopaths are in this way “into” or “behind” the mischief and pain is what constitutes their malice…they can set themselves directly against others’ aims as such. Part of their intent in cruelty is to suborn others’ agency to theirs, in a way that seeks the others’ recognition of their ability to do just this, like a child hurtfully teasing a dog or sitting on her little brother. (ibid., 316)

It is precisely because of this complexity that psychopaths are responsible in the sense of “attributability.” And this is sufficient to warrant hard feelings and the confrontation of psychopaths as antagonists. At the same time, the full range of responses to responsible agents who are also accountable is simply not warranted. Thus, “resentment and indignation proper” are not appropriate in such cases.

This subtle picture has great explanatory power when it comes to our puzzling reactions to psychopaths. It can explain why on the one hand we blame, and on the other we think that they lack the very capacities required for justified blame. We are right to blame in one way, but not in another. We appropriately regard psychopaths as responsible in the attributability sense or face, but would make a mistake if we were to hold them accountable for their actions. And conceptually, the two sorts of moral appraisal appear distinct.

3. The Unified Blameworthiness View

Despite the appeal of the picture that recognizes two distinct and separable kinds of moral appraisability, there is an alternative picture that incorporates a different view of the relationship between attributability and accountability. On this view, attributability is sufficient for
accountability. If someone is properly described as cruel, for example, then he is accountable for his actions. To see the appeal of this picture, it is helpful to think about other sorts of cases. And indeed, there appear to be, unfortunately, many other cases of apparently “incorrigible” behavior to explore. Consider the case of a slave-owner in the southern United States in the 18th century. Perhaps here, too, it is impossible to bring such a person around to see the sense of the demand to free his slaves. Even more plausible is the supposition that, at the time they acted wrongly, they lacked the capacity to have responded to the reasons for acting differently. Does the view require us to say that, as in the case of the psychopath, making demands of the slave-owner is senseless and thus, that he is not accountable? After all, it would seem eminently reasonable for the slaves to resent his actions and hold him accountable. So treating the slave-owner as the Middle Course View treats the psychopath seems counterintuitive.

This is a serious challenge to the picture that distinguishes so clearly between two faces of responsibility, and to the related solution to the puzzle of the psychopath by appealing to that very distinction. It is worth following out an alternative picture that preserves this view of the slave-owners—and white supremacists and Nazis and Mafiosi—as deserving of the full panoply of blaming emotions. In particular, one might think that what these sorts of cases really reveal is that resentment and indignation are appropriate just so long as someone has unjustifiably exhibited the sort of disrespect and attitude of being “against” one that Watson describes in the case of the psychopath. We are making a fine distinction when we say to the victim of the psychopath, and equally to the slave, that hard feelings are appropriate, but resentment and indignation are not. In that case, it would seem that the psychopath, no less than the incorrigible slave-owner, is responsible in all the faces of responsibility, including that which warrants the reactive attitudes.
This last piece of reasoning about the slave-owner and related cases suggests a different solution to our puzzle, namely, that despite the psychopaths’ lack of capacities, it is fully appropriate to blame them, to make demands, and to react with the reactive attitudes of resentment and indignation. It is enough, on this view, that they are rationally competent. That is, they can recognize reasons for doing things and for refraining from doing so, even if they lack the ability to recognize moral reasons. This ability is sufficient for agents to show disregard to others, by taking their needs and desires not to be reasons for action, and, in turn, this is sufficient for the kind of bad quality of will that warrants resentment and indignation.

Recently there have been several subtle defenses of this view, inspired by that of T.M. Scanlon, including those of Talbert and Smith.\textsuperscript{15} It will be helpful to set out the premises in an extended argument applying this view to the psychopath.

The argument from psychopaths’ cruelty to accountability

(1) Psychopaths’ actions manifest contempt or cruelty towards others.

(2) If one’s actions manifest contempt or cruelty, one is worthy of resentment.

Therefore,

(3) Psychopaths are worthy of resentment. (1, 2)

(4) If one is worthy of resentment, then one is accountable.

(5) Psychopaths are accountable for their actions. (3, 4)\textsuperscript{16}

Evaluating the key premises of this argument will allow us to explore both an assumption common to both pictures and to home in on the heart of the disagreement between them.
4. Questioning a Shared Commitment Regarding the Psychopath

Premise (1) would seem to be uncontroversial. After all, even Watson accepts it as the claim underlying his acceptance of the conclusion that psychopaths are responsible in the attributability face. Thus, it is a shared assumption by advocates of both approaches to the psychopath described thus far. But what reason is there for accepting it? After all, one might object that psychopaths, lacking moral understanding and the ability to take others’ interests as reasons for action or refraining, cannot express contempt or ill will or cruelty through their actions. Taking this objection seriously, Talbert, for one, takes this to be something that needs arguing for, and he obliges.17

Talbert acknowledges that the psychopath cannot make a judgment with moral content of the sort, “Your moral standing is not a reason for my refraining,” but he can make other judgments that do show ill will. Because the psychopath can understand reasons for acting and refraining (e.g., “I feel like a beer, that is a reason to go rob the 7-11”) he can make judgments like the following: “Your pain is no reason for me not to act.” Or he can simply deny that there is any reason at all not to injure you (“Blame and Responsiveness,” 522). Such judgments convey disrespect and contempt.

But there is a considerable gap between lack of respect on the one hand and disrespect on the other. To see this more clearly, consider an example from Watson (who, we have seen, endorses (1), but would reject this form of reasoning to it). The Psychlops is “a simple creature with two basic ends: to eat food that it finds tasty, and to preserve its own life.” It can make judgments such as “F’s flavor counts in favor of eating it,” and “if doing x would lead to eating something tasty, that counts in favor of doing it.” Perhaps these judgments even “implicate” the
judgment that the lethal effects of eating a human being do not matter or count as a reason for refraining. Still, it is to go too far to conclude that such a being exhibits contempt or cruelty.

Because of this, Watson offers a different kind of reasoning to (1), as we saw above. Psychopaths, being more complex, can “set themselves against us.” Some, at least, actively enjoy causing pain. It is this set of qualities that licenses the conclusion that they show contempt or ill will.

I believe there is room to question even this reasoning to (1). Again, it must be emphasized that we are assuming the features stipulated of the psychopath. This is consistent with its being the case that actual psychopaths really are cruel and contemptuous. Still, sticking with the stipulations, and assuming that there really is no moral understanding, and no capacity to see others’ interests as reason-giving in any way, it is not clear that they really can exhibit more ill will than the Psychlops. They may be more dangerous (because our pain is one of their ends, and not one among possible other means, or byproducts of means, as in the case of the Psychlops). And we may have more visceral reactions precisely for that reason. But the question at hand is whether their pleasure in our pain entails contempt or disrespect, or even cruelty as we understand it when we blame people for it.

To begin to see how we can question this seemingly obvious entailment, consider the opposite sort of case. Imagine a creature who lacks all moral understanding, and does not see others’ interests as fundamentally reason-giving, but who, along with beer and cigarettes enjoys watching car chases and other people enjoying themselves. He doesn’t care about other people in the sense of taking their interests to be reasons for acting. He just enjoys seeing them having a good time. In such a case, in my view, we would be reluctant to describe the case as one in
which the creature is kind. Similarly, it is not clear why we should describe the psychopath as one who is cruel.

In response, one might add that in the psychopath’s case, there is an enjoyment of the agency, and even power, in creating the pain. Yet even here, it seems possible to create a parallel for pleasure. Just as one might prefer to be the agent of success in a video game than watch someone else do so, the psychopath we are imagining might enjoy being the agent of another’s pleasure more than simply watching it. These things don’t add up to kindness unless one does something for the sake of someone. I am not exactly sure what the parallel of doing something for the sake of someone is in the case of cruelty or acts manifesting contempt. But the mere taking of pleasure in the pain, and even in being the instrument of pain does not entail cruelty or contempt. The addition of any of a variety of other judgments would turn the case into one of cruelty or contempt--including, but not limited to, ones such as “his moral standing is not a good enough reason to refrain.” At the very least, considering the kindness case shifts the burden in the following way: unless there is an asymmetry in what is needed for kindness and cruelty when it comes to moral understanding, the mere taking of pleasure in being the instrument of others’ pain is not by itself sufficient for cruelty. Call this “the symmetry argument” for the claim that psychopaths--understood as stipulated--do not display cruelty.

One might at this point object that while thinking about what kindness requires shows us that cruelty requires moral understanding, other moral vices, such as insensitivity to suffering, do not, and we can certainly attribute these vices to the psychopath. In this way, one might take issue with the idea that there is a general symmetry between moral virtues and vices when it comes to moral understanding. Yet, on reflection, I believe that we can identify several different senses of “insensitivity” to suffering, and while psychopaths as we are thinking of them are
surely insensitive in some senses, they are not in others. In one sense of sensitivity, one is sensitive to suffering when one is capable of correctly categorizing those in pain and those who are not, say. Psychopaths seem sensitive in this sense. But this kind of sensitivity is consistent with moral vice and virtue: one can respond in morally appropriate or morally inappropriate ways to suffering. In a second sense of insensitivity, one might be unable to recognize moral facts such as that someone’s being in pain provides a reason to discontinue one’s behavior that causes her pain, for example. And in a third sense of insensitivity, one might be capable of grasping these moral facts, but simply not care about them or let them affect their deliberation. Given the characterization of the psychopath at hand, the psychopaths are indeed insensitive in the second sense, but this kind of insensitivity is consistent with a lack of moral vice. To see this, consider again the Psychlops who would rightly be described as insensitive in this sense. It seems to be the third kind of insensitivity, a kind of motivational insensitivity, that constitutes a moral vice. This kind of insensitivity really amounts to a kind of indifference; one can grasp certain kinds of moral facts, but either one does not care to make the effort on given occasions, or one does not care to act on one’s recognition when one does. I believe that at least sometimes when insensitivity is claimed to be a moral vice, this is because various forms of insensitivity are not clearly distinguished, and the default background assumption is that people have the capacity to grasp moral facts. Psychopaths—at least as characterized for present purposes—provide a useful case precisely for allowing us to see that traits that so often go together can come apart.\textsuperscript{18}

Now it may very well be that in all real cases, the taking of pleasure in others’ pain co-exists with sufficient understanding of the concept of the moral standing of others. It is also worth noting that the conclusion of the symmetry argument is that a variety of judgments of moral qualities are incorrectly made of the psychopath. This leaves open that other attributions
are correctly made, perhaps even including some that can have moral associations in other contexts, such as a most general form of insensitivity that includes psychopaths, the Psychlops, and incorrigible racists—that is, a general failure to respond, for whatever reason, to suffering of others in one’s behavior. Such a trait, when combined with a general capacity for moral understanding, can turn into the kind of indifference that is only possible when one is aware of something and still does not care. Further, psychopaths can be properly described as committing themselves to ends. Thus, insofar as there is a non-moral variety of responsibility in the attributability sense, psychopaths can be subjects of the attribution of non-moral faults. At the same time, in the moral variety that I take to be at issue here, they will not be responsible even in the attributability sense if the symmetry argument is sound. The distinction between the moral and non-moral is a subtle one.\textsuperscript{19} Yet thinking about the cases, particularly the positive ones, helps to reveal something important about the nature of moral virtues and vices. It does so by making clear the connection between what appear to be genuinely moral virtues (and, indirectly, vices) and moral understanding. For we have an intuitive grasp that insensitivity (understood as not picking up on salient moral facts such as that someone’s suffering is a bad thing) and indifference (understood as picking up on them and not caring) are two different kinds of faults. And this, together with an explanation in terms of the presence or absence of moral understanding, can help us in turn to see how cruelty and the kind of defect of the agent who merely takes pleasure in causing others pain are also different in kind.\textsuperscript{20}

Now it might be asked at this point what hinges on the distinction between moral traits, such as cruelty on the one hand, and non-moral ones, such as insensitivity in the sense of not picking up on the badness of others’ pain and the (mere) enjoyment of causing pain on the other hand, particularly if it is agreed that we respond negatively to both sorts. The answer is in the
details. As we have seen, there are many different kinds of negative responses, including resentment and hard feelings, and it might turn out that some are appropriate in the first kind of case but not in the second. I return to this point in section 7. For now, it is important to note that the upshot of the symmetry argument, then, is that in the hypothetical case in which agents take pleasure in causing others pain without sufficient understanding of the concept of the moral standing of others, agents can have a number of bad qualities, but not cruelty and other paradigm moral vices in particular.

A final objection to the symmetry argument targets one its key presuppositions. The idea is that two things assumed to come apart—namely, an understanding of pain and moral understanding—can’t in fact come apart. Indeed, one might argue that the very understanding of pain—and pleasure—requires an understanding of others’ interests as reason-giving. One might reason as follows: to understand pain, one must understand that it is bad for the person who experiences it, and in this way necessarily provides a reason not to cause it. This is an important objection, and I believe the best reply for our purposes is to modify the reasoning just offered for the claim that psychopaths are not cruel and morally blameworthy in the attributability sense. We can acknowledge the possibility of this robust conception of pain by transforming the argument into a dilemma: either the understanding of pain and the understanding of others’ interests as reason-giving come apart, in which case the symmetry argument as already presented succeeds in casting doubt on the claim that psychopaths (as supposed) are cruel and blameworthy in the attributability sense, or they do not come apart, in which case we cannot make the supposition we are asked to make about psychopaths in the first place.

If the first horn of the dilemma is correct, then the taking of pleasure in others’ pain is sufficient for our regarding such beings as particularly dangerous. But we will have good reason
to doubt that this is sufficient for the sort of contempt or cruelty that is claimed to justify the reactive attitudes.\textsuperscript{21}

On the second horn of the dilemma, it would follow that psychopathy could not be used in the way it has been to illustrate a distinction between the two faces of responsibility in the first place.\textsuperscript{22} But it would not yet follow that attributability and accountability cannot come apart. To evaluate this further claim, we will need to evaluate the rest of the argument for the Unified Blameworthiness account.

Either way, we will have reasons to question (1)--either because our initial suppositions can’t be coherently made, or because we have an argument against it, when we do make the suppositions. Having questioned (1), I have departed from the way both pictures have been applied to the psychopath in particular. But the difference between the pictures still has great importance. Let us turn then to premise (2), which captures the central difference in approaches between the Middle Course View and the Unified Blameworthiness View. Where (1) is a claim about the nature of cruelty and its relationship to our stipulated conditions for psychopathy, (2) is a claim directly linking cruelty to a key aspect of accountability. Now that we have a fuller picture of cruelty itself, we can turn to an evaluation of the claim that it is sufficient for the appropriateness of the blaming attitudes of accountability.

5. A Defense of the Insufficiency of Attributability for Accountability, and the Role of Demands

One key aspect of the Middle Course View is that contempt and cruelty suffice for attributability, but not accountability. Since, on this picture, accountability is a matter of demands, and making demands is a kind of laying down of claims to adverse treatment and
reactive attitudes that dispose one toward such adverse treatment, it will not be enough for resentment that someone shows contempt or cruelty. Now I think that there are at least two ways that we can elaborate on this reasoning. One focuses on the idea that making demands presupposes “the capacity to understand moral address” (call this the “Argument from the Intelligibility of Moral Address,” or the “Argument from Intelligibility” for short.) The second focuses on the idea that making demands presupposes the capacity to act in particular ways (call this “Argument from Capacities to Meet Moral Demands” or the “Argument from Capacities” for short). In this section, I lay out both of these lines of reasoning, emphasizing at least one advantage for the second, and in the following sections, I consider arguments on behalf of the Unified Blameworthiness View that challenge both lines of reasoning.

The first line of reasoning, spelled out by Gary Watson, is one that appeals to a kind of “conceptual inaptness” of making a demand of “a creature that is incapable of recognizing one’s standing to make demands.” Because such a creature is “incapable of the reciprocity that demanding and owing justification presumes, moral criticism is not only futile, but senseless” (314). Watson makes very clear that he is not claiming that making demands is unfair; rather, doing so is inapt because such creatures do not meet the conditions for intelligible moral address. “Nothing that they could do could be intelligibly construed as an apology or acknowledgement.”

The reasoning can be represented as follows:

The Argument from the Intelligibility of Moral Address

(i) Resentment of an agent for an action is appropriate only if the agent is accountable for that action.
(ii) An agent is accountable for an action only if it is apt to make demands of that agent.

(iii) Resentment of an agent is appropriate only if it is apt to make demands of the agent. (i and ii)

(iv) One can aptly make demands of an agent only if the agent has the capacity to comprehend moral demands as normatively relevant in the right way.

(v) An agent can act in a way that shows cruelty and at the same time lack the capacity to comprehend moral demands as normatively relevant in the right way.

(vi) An agent can act in a way that shows cruelty without resentment of the agent being appropriate for that action. (iii, iv, and v)

The psychopath is meant to illustrate (v). But if the symmetry argument is correct, then it is not obvious that (v) is true. Cruelty may require a kind of moral understanding after all. In other words, it may require an understanding that others’ pain provides some reason not to cause it. In that case, the psychopath, as defined here, does not clearly illustrate (v), then. And the same considerations raise the worry that (v) will be difficult, if not impossible, to sustain. I believe that there are ways--at least initially--of resisting this objection. For example, it might be possible to distinguish between different kinds, or levels, of moral understanding, so that one can have the kind required for cruelty--including understanding that others’ pain is bad for them and thus, reason-giving--but not the kind required to understand the relevant demands in the right way. But while I ultimately agree with the conclusion that one can be cruel without being accountable, I believe that another line of argument that can avoid this particular challenge concerning moral understanding is promising. In the rest of this section, then, I turn to the
second line of argument. The fundamental idea is that demands are futile if the one to whom they are directed cannot meet them.

After all, it makes no sense to demand that you fly to Mars this afternoon; we both know you can’t do it. But since resentment and the other blaming attitudes associated with accountability are appropriate only if it is apt to make relevant demands of their targets, then if the relevant moral demands are, to the psychopath (or to anyone else), like demanding that you fly to Mars, then resentment could not possibly be appropriate. The reasoning goes as follows:

The Argument From Capacities to Meet Moral Demands

(a) Resentment of an agent for an action is appropriate only if the agent is accountable for that action.

(b) An agent is accountable for an action only if it is apt to make demands of that agent.

Therefore,

(c) Resentment of an agent is appropriate only if it is apt to make demands of the agent. (a, b)

(d) It is apt to make demands of an agent only if the agent has the capacities needed to meet the demands.

(e) An agent can show cruelty without having the capacities needed to meet the demands.

Therefore,

(f) An agent can show cruelty without resentment of that agent being appropriate. (c-e)

At this point, some clarification and elaboration is essential. First, it is important to see that the two arguments are similar in structure, and begin with the same conception of accountability. Where they differ is in what each takes to be required for demands to be apt. The Argument
Intelligibility requires that the target of appropriate resentment find moral demands intelligible, where the Argument from Capacities requires that the targets of appropriate resentment be able to meet the demands. It is an open question at this point that both arguments are sound, and even that the requirement that one have the capacities to meet demands in the end can be shown to entail the requirement of intelligibility. The arguments do have different implications, however, in the scope of agents they rule out as accountable agents insofar as it is possible for an agent to find moral address intelligible and yet lack the capacities needed to act on the relevant demands.

This brings us to two ways in which the Argument from Capacities requires elaboration. First, it is important to clarify further what the contents of the relevant moral demands are, and, second, it is important to expand on the notion of capacities at issue.

Let us take the question of demands first. One might initially ask whether the Mars case is a poor analogy because while it is obvious that no one can fly to Mars this afternoon, it is not at all obvious that psychopaths (or anyone else with enough sophistication to rightly be called cruel) lacks the ability to comply with the relevant demands to act differently than they in fact do. After all, just about everyone under consideration here—psychopaths, incorrigible racists, and others whom we might even consider calling cruel—can meet demands of some kinds, such as demands to stop causing pain. They can stop doing what they are doing, whatever their specifically moral capacities are like. As Watson points out, they can be influenced by threats, for example, without having any moral understanding at all.

If in fact the demands in question are simply demands to perform an act of a different act-type, where the relevant act-types include putting the gun down or moving away, say, then this line of reasoning would not be off to a good start. However, there is good reason to think that
the demands in question are not of this type. Rather, the demands in question are demands to act in such ways for certain reasons. Now I think that there are different ways that one might support this claim about the content of the relevant demands. Here is one: Insofar as demands are understood as a way we hold accountable, we need to be able to distinguish them from mere threats or even offers, for example. And the most natural way to do that is to take it that what we are aiming at in making our demands is that one act for certain reasons. Thus, the Argument from Capacities is best understood as one that takes moral demands to be in part about agents’ reasons for acting.

A second question concerning how we should understand the argument centers on the notion of capacity invoked in (d). For we could understand capacity here in either a narrow or a broad way, where to understand it narrowly would be to understand it as indexed to the time of action. Does one, at the time one acts badly, say, have the capacity to have met a demand to act otherwise? If so, then one has the capacity in the narrow sense. Alternatively, one might have the capacity in the sense that one could have earlier acted differently (with foresight) so that one would now have the capacity in the narrow sense, but one does not have that capacity in the narrow sense at the time of acting badly. How we answer this question has serious implications about the scope of the argument. If we opt for a broad reading of the requirement that one have the capacity in order for it to be apt to make demands, then it might turn out that many who appear to lack the capacity to meet them at the time they act badly are nevertheless still open to resentment and to being rightly held accountable. Perhaps some racists, for example, would naturally fall into this category. Perhaps they are not now able to meet the relevant demands, but that is only because they acted with foresight earlier in ways that served to cultivate the very deficiencies of understanding that now constrain them. In this sort of case, there is something
appealing in the idea that one does not thereby get off the hook merely because one earlier incapacitated oneself. While the issue is a subtle one, for now, I want to leave open the possibility that one can be accountable for a specific action even if one cannot at the time of action meet the relevant demands in such a situation. Thus, let us understand the requirement of capacities for the aptness of making demands to be disjunctive: either one has the capacity in the narrow sense, or one’s lack of a capacity at the time of action is one for which one is culpable. I return to questions about capacity in section 6. With this elaboration of the argument, we can see that in addition to sharing an understanding of accountability with the Unified Blameworthiness View, the Argument from Capacities hinges on a plausible premise concerning the conditions under which making demands is apt, and a plausible premise concerning the conditions under which one might lack the relevant capacities to make those demands apt.

6. Meeting Challenges from the Unified Blameworthiness View: The Content of Demands

At the same time, the Argument from Capacities, together with the Argument from Intelligibility, still faces serious challenges from the Unified Blameworthiness View. In this section and the next, I consider two independent challenges to the conclusion of those arguments, and in favor of the sufficiency of blameworthiness in the attributability sense for blameworthiness in the accountability sense.

The idea behind the first challenge is that if one’s actions manifest judgments of disrespect, then whether or not one has voluntary control over one’s actions (or attitudes for that matter), it can still be apt to make the kinds of demands that are part and parcel of our accountability practices. Thus, the reasoning goes, the link between the expression of contempt
or cruelty and the appropriateness of making demands remains unbroken. (As applied directly to the Argument from Capacities, this reasoning can be seen to challenge premise (d).)

Smith (2008) offers a subtle defense of this idea. Unlike mere appraisal of objects or people as good or bad, athletic or not, handsome or not, moral appraisal already brings with it a distinctive quality that undergirds our reactive attitudes. As Smith writes,

Moral criticism, by its very nature, seems to address a demand to its target. It calls upon the agent to explain or justify her rational activity in some area, and to acknowledge fault if such a justification cannot be provided. If I am called ugly, I may feel embarrassed, amused, or incredulous, but what I am not likely to feel is challenged: that is, I will not see any reason to try to defend myself against the charge (or to apologize or make excuses for my alleged unsightliness). If I am called selfish, by contrast, I cannot regard that as merely an unwelcome description of some trait or feature I possess: it is a direct challenge to me as a moral agent, and therefore implicitly calls for some kind of response on my part…

I believe it is this feature of moral criticism--that it addresses a justificatory demand to its target--that constitutes the relevant “reactive entitlement” that Watson regards as essential to our practices of moral accountability. (“Control, Responsibility, and Moral Assessment,” 381)

On the account Smith is defending and elaborating here, there is really one notion of responsibility, and the conditions for it are a kind of self-expression in the judgments manifested in one’s action. If one displays in one’s actions the judgment that someone else’s pain is not a
reason to refrain from an action that will hurt her, this is sufficient for demanding a justification or reassessment of this very judgment. And the making of such a demand is continuous with other reactive entitlements such as resentment and indignation. All are further reactions licensed by the reasons-sensitive attitudes expressed in the actions, and these additional reactions all go beyond mere appraisal.

I believe that this line of reasoning illuminates something important, but that it is not reason to give up a distinction between attributability and accountability. At least, I would argue that on the best understanding of the demands relevant to accountability, the demands in question are not the same as the ones appealed to by the Unified Blameworthiness account. A demand to justify one’s actions and underlying judgments is one thing; a demand to act differently than one does is another, or to act for different reasons is another. Smith’s account takes demands to be for justification, while the view that takes accountability and attributability to come apart, as I have understood it here, takes the relevant demands to at least include the demand to act differently or for different reasons.

In this way, the relevant demands are, arguably, different from demands to provide justification. One might argue that demands to provide justification also require the real capacities needed to meet those demands for them to make sense. But many take it that one can be epistemically irrational without having an ability to do otherwise; that is, it can make sense to criticize someone for holding unsupported beliefs without thinking that they have control over whether they do, or over whether they could have voluntarily avoided doing so. Similarly, one might argue that it makes sense to demand that people justify their judgments and actions, even when they cannot. In contrast, the demand to act differently might be thought to be different precisely because it invokes a kind of “Ought Implies Can” principle. If demands that someone
act differently than she does, or with different reasons, are just, these presuppose that she can do so. Thus, if it is possible to show contempt in one’s actions, but lack any relevant capacity to do otherwise, then we would indeed have a case in which demands of the kind that most clearly invoke something akin to the Ought Implies Can principle would not make sense.30

Is there a reason to prefer one conception of the relevant demands over the other? One reason one might initially think the move to justificatory demands is attractive is precisely that it does not require agents’ ability to do otherwise in order to be blameworthy, and sidesteps the need to invoke anything like the Ought Implies Can principle. But it is not clear that the move to justificatory demands avoids such a commitment in the end. To see this, consider a challenge Smith acknowledges and her response to it. She writes, “Can it be appropriate to address a ‘demand’ to an agent to justify something she has not deliberately chosen, and over which she may not have direct voluntary control? Does this not violate the established “ought implies can” principle?” (ibid., 382). In response, Smith notes that “these are large and important questions,” and goes on to offer us “as a start” a reminder of “how often we do, as a matter of fact, both make and accept demands to justify things that we have not deliberately chosen, and which may not be under our direct voluntary control.” For example, the following demands are familiar, at least in their type: “How could you have forgotten our anniversary?” “Didn’t you realize how insensitive that was?” “How could you have left our child unattended at the park?” and so on (ibid., 382). Further, Smith observes, it seems that in these cases the object of our criticism is not a person’s voluntary choice, but the person’s judgment that is implicit in her attitude (ibid., 383). The conclusion seems to be that once we move to a picture in which the relevant demands are justificatory ones, observation of our actual practices will suggest that we do not in fact require that one could have acted differently when one acted.
Initially, as we saw, it might seem that one could avoid worries stemming from violations of the Ought Implies Can principle simply by shifting from a demand to act in a certain way to a justificatory demand. But a variant of the challenge still remains. Insofar as demands are apt, it seems that they require that one ought--and so, can--fulfill them. And, in fact, the demand to justify or apologize is a demand for something beyond the capacity of the psychopath on the characterization we have been working with. Smith’s initial posing of the challenge has it that the alleged violation of Ought Implies Can would happen if the agent couldn’t have acted differently in the first place; according to what I see as the more direct challenge for her alternative suggestion is that the demands would be violated if the agent can’t now act to justify (or apologize for) her earlier action.

This challenge is important, among other reasons, because on Smith’s own understanding of a cruel person as “someone who judges that the fact that something will cause pain or suffering to another is no reason to avoid it (and is perhaps even a reason to pursue it)”, a cruel person could simply lack the capacity to comply with the demand for justification and apology as easily as he could lack the capacity to have failed to respond to the relevant reasons at the time of his cruel behavior. Thus, unless this challenge is answered, the move to understanding the relevant demands as justificatory will not help to undermine the distinction between attributability and accountability.

There is a second challenge faced by this view, as well, that is quite independent of the first. The challenge is to the very idea that blameworthiness in the accountability sense is captured by the condition that such justificatory demands are appropriate.

It is true that the examples of demands that Smith offers sound very much like instances of blame. But it is not clear that they are pure demands for justification. It is possible to respond
to a given piece of behavior by genuinely asking for one of justification or apology or correction of a mistaken belief, without taking this to be a form of blame at all. I might assume that someone is answerable, but depending on how they actually answer, I might form a variety of different moral judgments about them, including that they were blameless. A mere demand for justification is not blame; to turn into blame, such a demand must already presuppose a judgment of blameworthiness. And similarly a judgment of blameworthiness would seem to presuppose, at the least, that there is no justification.

Further, if the conditions of accountability can be satisfied simply by the making of certain kinds of judgments, then it appears that the concept of excuse—as distinct from justification—is not relevant. It is unclear what could count as an excuse or mitigating factor in blameworthiness on this model. Assuming that one’s action and the attitudes revealed in the action are not justified, then it would appear that there would be no further work for excuse to do on this model. Suppose, for example, that someone under great stress, or a young child, acts in such a way as to reveal an unjustified judgment that someone else’s interests do not matter, or that they do not matter very much, in a given case, when they do. The fact that the agent is under stress or is a child does not affect the agent’s blameworthiness on this view. Alternatively, if justification is used in a broader way than is typically done to include excuse, then it would seem that there is some additional factor, other than the judgment itself revealed in action, that is the basis of the moral appraisal. If one might be excused for acting on a judgment that reveals cruelty or a failure to prioritize another’s feelings, for example, then something other than the judgment itself—perhaps something relating to one’s circumstances or one’s abilities—is relevant to the ultimate judgment of blameworthiness after all. One might reject the idea that there are
excuses, but to do so would be highly revisionary, and it would be essential to have an explanation for what would seem to be a systematic error.

In sum, the shift to conceiving of demands as justificatory captures something of significance, and appears initially to be a promising way of rejecting the Argument from Capacities in particular. But the strategy faces two serious challenges. The first is that even if we grant that the relevant demands are justificatory in nature, it is not clear that this allows for a side-stepping of the challenge that demands are appropriate only when their targets have the capacities to meet them. Thus, it is not clear that even the move to justificatory demands will by itself undermine the distinction between attributability and accountability. The second is that the picture of accountability as the aptness of such justificatory demands, and in turn, the understanding of the aptness in terms of their targets having manifested judgments in their actions fails to accommodate important phenomena. Thus, though Smith’s observations about our practices are important and suggestive, I do not believe that they give us sufficient reason to abandon a picture that takes the demands relevant to accountability to include demands to act differently.

Arguably, the incorrigible racists, the 18th century slaveowner, and certain Nazis and White Supremacists, are beings of whom all of these things are true: their actions show clear contempt, cruelty, and disrespect, and yet the making of demands that they act differently in the first place does not make sense in the sense that they cannot be complied with. They can make the kinds of judgments that the psychopaths cannot, namely, that while certain people have interests, these interests do not provide sufficient reason not to harm them for sport, for example. (It may make sense to appear to be making the demands for a variety of reasons, but in the sense of genuinely demanding that they do something that they lack the capacity to do, it does not
make sense to demand.) If this is correct, then it leaves open the possibility that there really is a clear case of attributability without accountability, contrary to the Unified Blameworthiness view. But contrary to Watson’s categorization, the psychopath does not instantiate this and the incorrigible racist does.

7. Meeting Challenges from the Unified Blameworthiness View: The Point of Demands and the Reactive Attitudes

It remains to consider one more argument on the other side; an argument for the claim that attributability is sufficient for accountability. I think that it can be developed in a variety of ways, so that it may be seen to target more than one premise in the Argument from Capacities. The idea behind it is simple and powerful and appeals to our intuitions when asked to judge the responses of victims of horrible wrongdoing.

Talbert has elegantly developed this sort of reasoning. As he writes: “It seems to me that blame is not without a point just because it fails to initiate a moral dialogue and reform. Such outcomes can be an important goal and result of blame, but we need not say that blame is inapt when these outcomes are unachieved or unachievable” (“Blame and Responsiveness,” 106). Rather, at least one point of blame is our own expression of protest and standing up against the judgments that have been expressed in the actions of those who have harmed us: “Even when it is impossible to inspire consideration of our moral standing in others, insisting on this standing can still serve to affirm for ourselves certain moral facts, and this can be something of significant moral value” (ibid., 106). The slave or the victim of white racism is not doing something pointless in expressing an attitude of resentment or indignation, for example.
Now I think that there are two different ways to understand this objection. On one reading of Talbert’s reasoning, protest incorporates demands, and thus, it is intelligible even if the target cannot meet them. (This would then be another way to question (d) in the Argument from Capacities, and (iii) in the Argument from Intelligibility.) While Talbert acknowledges that the insistence on the part of the slave that his moral standing be recognized will necessarily miss its target in such a case, it does not follow that the insistence itself is “practically infelicitous.” Here, I think that we simply find two different notions of “aptness” or “intelligibility”. It seems to me that one can question (d) on these grounds only if one equivocates on the notion of “aptness.” I take it that the sense in which the demand must be apt for accountability is for it to be the case that it hits its target, so to speak. In contrast, the way in which the making of the demand can be apt that is deployed here has its focus on the victim rather than on the agent in question. And while the victim’s response is unquestionably important, there is reason to think that the way that demands must make sense for one to be accountable is a matter that has its focus on the agent. This is somewhat metaphorical, but accountability is a matter of demands in the sense that they have a point with respect to the very person who is accountable. The making of demands may have other undoubtedly important points, such as standing up for oneself and sending a message that one is on the side of right. This is one way that making demands can be apt. It might even be, for example, that making demands have some other point in a particular circumstance--e.g., your doing so impresses your boss and you get the job in the lab that allows you to find the cure for Alzheimer’s disease. But this is not related to the target of the demand in the right way to make the supposed target of the demand accountable.

Still, there is another way to develop the initial intuition that resentment on the part of the slave is appropriate. The idea is that resentment--and protest, and insistence--are in place
whether or not there is a demand made at all. This would be to question (a) or (b) in the Argument from Capacities. I take it that the notion of demands is really essential to accountability, and so (b) remains untouched. But one might question (a), the claimed connection between accountability and reactive attitudes like resentment. In doing so, however, it is important to note, one would be questioning an assumption shared by advocates of both the Middle Course View and the Unified Blameworthiness View. For (a) just is premise 4 in the argument that psychopaths are not only responsible in the attributability sense, but also in the accountability sense. So to conclude our evaluation of the debate, let us turn to that premise.

Could resentment—or at least a kind of resentment—be appropriate without accountability? If resentment is understood as disposing one to treating people in unwelcome ways, one might think that the burden is shifted onto the one claiming resentment is appropriate in such a case. For it might seem that one is only entitled to such attitudes and to treating people in adverse ways if one is entitled to demand that they act otherwise than they do.

But even here, there is an ambiguity. Even if someone is not responsible in any way, but is simply dangerous, then one might be licensed to treat that person (or animal) in ways that happen to be unwelcome. One distances oneself, for example. This is a sense of the appropriateness of adverse treatment that is different, however, from a notion of deserved sanction that one might think is built into the idea of accountability. If there is a notion of resentment that does not presuppose that adverse treatment is deserved, but rather, is a kind of protest that may have as a result unwelcome treatment, then resentment is not limited to accountability, at least not understood in a way that links it to deserved sanction. This may be a kind of resentment that marks an important attitude and moral practice. Notably, this reasoning would lead to a rejection of the unqualified claim that resentment is inappropriate for those who
are responsible only in the attributability face. However, it would not necessarily lead to a rejection of a view that recognizes an important distinction between two notions associated with responsibility. What it would show is that a modification of sorts to the picture is necessary. Some of the features associated with accountability— notably, demands and reactive attitudes— might themselves come apart in interesting ways, and on the view of resentment just developed, it would turn out that accountability is not the exclusive locus of all of the reactive attitudes.  

8. The Separable No Blameworthiness View: Applying the Picture

I have adopted insights from each of the two attempted solutions to the problem of the psychopaths that I have explored in depth, but I have also rejected aspects of each. Notably, I have argued that we ought to recognize a real distinction between attributability and accountability that is not merely conceptual, but that is also illustrated by cases in which aretaic judgments are properly made, but in which it is not appropriate to hold the agents in question accountable. At the same time, while I have committed to accountability as distinct from attributability precisely because of the aptness (and justness) of demands and associated obligations, I have not committed to demands in turn entailing deserved harm or sanction, or even to all reactive attitudes requiring accountability for their appropriateness conditions. And I have focused on demands that apply to agents at the time of blameworthy action, whether expressed or not.

The focus on demands alone, though, provides a sufficient basis for categorizing the incorrigible racists differently from the way advocates of both views have so far done. Sharing the inability to meet demands is something such racists—if truly incorrigible—have in common with psychopaths, and as a result, neither are morally accountable.
Further, attention to the stipulated deficits of the psychopath and to the robust requirements for attributability itself leads me to categorize psychopaths differently from the way advocates of both the Middle Course and Unified Blameworthiness Views have done, as well. Psychopaths do not obviously meet the minimal conditions for displaying attitudes of contempt and cruelty that make them responsible in this sense—as opposed to dangerous and dislikeable. In this way, psychopaths and incorrigible racists are very different from each other, the former not being morally responsible in either sense, while the incorrigible racists are responsible in the attributability sense.

Because the Middle Course view with which we began also recognizes a difference between the psychopath and the incorrigible racist (as we have been conceiving of both), it is worth explicitly contrasting the approach I have taken here and the Middle Course View. Watson, for one, sees the two cases as categorically different. Yet for him the distinction is between the psychopath who is responsible only in the attributability sense and the incorrigible racist being responsible in the accountability sense, whereas on my view the psychopath is not responsible in either sense, and the incorrigible racist is responsible only in the attributability sense. According to Watson, the more ordinary apparently incorrigible person (such as a Nazi or Mafioso or white supremacist) “occasionally makes a genuine return to the moral point of view…In contrast, I am supposing, there is nothing to which to return in psychopaths” (“Trouble,” 318). This seems to me to capture something important. But it is incomplete. On one way of developing this point, the fact that sometimes slave-owners make such a return is relevant because it points to a possibility, or opportunity, that they have and that psychopaths lack. But this approach is not available, for by hypothesis the incorrigible slave-owner has no more opportunity than the psychopath. Further, given that Watson endorses the Argument from
Intelligibility, the point is naturally developed in a different way. Demands of a sort can still be apt even if the slave-owner cannot meet them, and that is because he has certain general moral concepts that make the demands at least potentially intelligible. (Importantly, this idea goes beyond the Argument from Intelligibility, because it makes intelligibility sufficient, and not merely necessary, for accountability.) But since psychopaths lack the very concepts needed to even begin to understand the demands as normatively relevant in the right way, the demands are inapt. On this reading, because demands can be apt without their targets having the capacity to meet them, it seems that something like the Ought Implies Can principle is rejected. Thus, this way of distinguishing the psychopath and incorrigible racist rules out appeal to the Argument from Capacities. On my account, in contrast, the important distinction between the psychopath and the incorrigible racist can be captured on the basis of the same difference in their deficits, but without rejecting the idea that apt demands require the ability to meet them.

At this point, one might resist the idea that any 18th century slave-owners really were incorrigible. It might even be easier in some ways to imagine a being without any moral knowledge than one who for some reason can’t be brought to see that some subset of people have moral standing and whose interests give sufficient reasons for acting in certain ways. They certainly have more in the way of the conceptual apparatus than psychopaths as we have defined them. And it is tempting to think that actual slave-owners could have seen—and should have seen—that their slaves had the same moral status as they themselves and acted accordingly. Perhaps they were operating on false empirical assumptions that they ought to have critically examined. Particularly if we understand the capacities required to meet the relevant demands in the broad sense, then it will be easier to conclude that such slave-owners had the relevant
capacities after all. Either way, it is possible that the case of the truly incorrigible slave-owner becomes merely hypothetical.

At the same time, it is possible to imagine that there are people, perhaps with various forms of insanity or mental impairment who simply cannot, in any sense, act differently than they do, but whose actions express the judgments that other people’s interests are not reason-giving, or not reason-giving to any significant extent.

And, as mentioned earlier, it is an open question whether psychopaths as defined by, say, the Hare Checklist, really are “morally blind” in the way that we have stipulated. Perhaps they are not, and some manifestation of their moral “sight” will turn out to explain and justify our actual practices of blaming those diagnosed as psychopaths in every sense, including holding them accountable. In that case, the psychopaths we have considered thus far also become hypothetical cases, but, I hope, no less interesting for that. Is there a way we can tell if there are in fact morally blind people?

Moral blindness is not itself a condition on the Hare Checklist. But it has been hypothesized to explain conditions on the checklist such as a lack of remorse, guilt, or empathy. Psychopaths’ poor performance on the Moral/Conventional Test (MCT), for example, in which they do not distinguish moral from conventional transgressions, has been much cited as evidence for this etiology (for example, see Blair, Mitchell and Blair, The Psychopath, 57-58). While those not diagnosed with psychopathy, and even relatively young children, count a case of breaking someone’s leg for fun as a moral transgression and a boy wearing a dress to school as a conventional one, psychopaths tend to count both the intuitively moral and the conventional ones as moral. Their failure to distinguish suggests that they do not really have a concept of moral reasons, but only one of rules that fail to make any distinctions between moral ones and others.
Recently, good questions have been raised about the MCT and what it shows. To take one example, Aharoni, Sinnott-Armstrong, and Kiehl, like Blair many years earlier, hypothesized that psychopaths’ over-inclusion of cases as moral might be explained by their desire to make a good impression on their testers.\textsuperscript{34} This hypothesis, however, is consistent both with an inability to correctly sort cases into moral and conventional categories (as Blair originally concluded) and with an ability to do so. So they devised a forced choice test, and subjects were told that exactly half the answers were moral and half conventional, thereby ruling out an incentive to please the testers by categorizing more cases as moral. They defined “moral wrongfulness” as “acts that society would consider wrong even if there were no rules, customs, or laws against them” (ibid., 5). Interestingly, they found that in this forced choice test, incarcerated psychopaths did no worse than their incarcerated peers on the task. The authors (among others) raise the possibility that moral blindness is not the root problem; rather, the problem is motivational: psychopaths do not care about moral rules.

But it would be hasty to draw any conclusions on the basis of positive results on sorting tasks such as this. While a true failure to sort correctly would point to a failure of moral knowledge, success in sorting does not entail moral understanding. To see this, consider that someone might sort moral rules from others because they learned all of the moral rules on Tuesdays and the other rules on other days, but this would not show the kind of moral understanding philosophers have thought necessary for acting for certain kinds of reasons. Thus, other sorts of tests, including those that ask subjects to offer explanations for why various actions are wrong, seem better suited to getting at the real distinction between the morally blind and, if it is possible, the merely morally unmotivated.\textsuperscript{35}
Although it is unsettled, much may ride on the question of whether psychopaths as diagnosed by the Hare Checklist and other measures are in fact morally blind. As long as there are genuine incapacities for moral action, demands will not make sense, and so psychopaths will not be responsible in the accountability sense. However, if the problem is not moral blindness, but moral motivation, then they are at least candidates for responsibility in the attributability sense. For, if such beings are possible, then, arguably, they would “stand for” the idea that the interests of others simply aren’t of sufficient importance to avoid harming them. Whatever the truth turns out to be about the psychopath, I hope to have brought out at least some ways in which imagining the possibilities can shed light on the very nature of moral responsibility.
NOTES

Previous versions of this paper were delivered at the University of Gothenburg in June 2012, the University of California, Santa Barbara in October 2012, the Mark L. Shapiro Graduate Student Conference at Brown University in November 2012, the Conference on Free Will at the University of Arizona in February 2013, the University of California, San Diego in April 2013, and the University of Tennessee, Knoxville in April 2013. I am indebted to audiences for their questions and discussion on each of these occasions. Special thanks also to Richard Arneson, David Brink, Gunnar Björnsson, Matthew Fulkerson, Matthew Hanser, Michael McKenna, David Palmer, Derk Pereboom, Sam Rickless, David Shoemaker, Matthew Talbert, Manuel Vargas, and Gary Watson for invaluable discussion of the issues and comments on previous drafts. Finally, I am very grateful to the editors of Ethics and two anonymous referees who provided extensive and incisive comments that greatly improved the paper.

1 Robert Hare, Without Conscience: The Disturbing World of the Psychopath Among Us (New York: Guilford Press, 1993): 33.


3 The labels for each of the views capture their respective commitments on (i) the question of whether attributability and accountability come apart (Separable or Unified) and on (ii) the question of whether or not the psychopaths are responsible in each sense (Middle Course for “yes” to attributability and “no” to accountability, Blameworthiness for “yes” to both, and No Blameworthiness for “no” to both). But as will become apparent, defenders of each view make additional commitments, including to distinctive conceptions of attributability and accountability
and to answers to the question of whether members of other groups are responsible in either sense. For ease of exposition, I have chosen to tie the labels to just these two key kinds of commitments.


5 (See, for example, James Blair, Derek Mitchell, and Karina Blair, *The Psychopath Among Us* (Malden: Blackwell Publishing, 2005).


9 One might object to the idea that making such judgments can be a kind of blaming, or attribution of responsibility at all. (Thanks to a referee for encouraging me to address this point.) Although I believe that it is natural to call making such negative judgments a kind of blaming, I also take it that the distinction between attributing moral fault and holding accountable can be described without calling both phenomena kinds of blaming or faces of responsibility. And the distinction can be used coherently to solve the puzzle of the psychopath without doing so in a way that is similar to the approach that recognizes two faces, or kinds, of blame. The solution would be to recognize that what might initially seem a type of blame is not blame, but instead an important type of moral criticism that is closely related to blame.
This sort of demand can nevertheless be related to blaming responses in a number of ways. First, demands that are expressed later, as part of blame itself, may depend in key ways on those to which the agent was subject at the time of action. For example, the same features of the target of blame may make both kinds of demands apt. Alternatively, it may be that blaming is a kind of holding to the demands to which the agent was subject at the time of action but did not fulfill. Sher, for example, suggests that blame can include a retrospective endorsing of the demand that applied to an agent at the time of action. (George Sher, “Kantian Fairness,” *Philosophical Issues* 15 (2005): 179-192. (p. 187).) Interestingly, Sher goes on to reject the claim that one must have the control required to meet a demand for it to be fair to make it, whereas in what follows I accept it.

In taking this to be central, I leave open whether demands must have the particular connection to sanctions and to reactive attitudes described by Watson. I take up this point briefly in section 7, and in more detail in Dana Kay Nelkin, “Desert, Fairness, and Resentment,” *Philosophical Explorations* 16 (2013): 117-132.

Watson sometimes suggests that the demands in question are demands to provide justification after the fact (see “Trouble.”) But as mentioned in the text, I believe that the more fundamental demands are those that apply at the time of action. I return to this point in section 5.

Whether actual slave owners were incapable of “being brought around” is not an easy question to answer. For now, let us suppose that there are—or at least can be—cases of this sort, and return to the empirical question later in the paper.

This kind of reasoning is nicely spelled out in the work of Talbert (“Blame and Responsiveness” and “Accountability, Aliens”), reasoning to which I return in section 7.

See T.M. Scanlon, *Moral Dimensions: Permissibility, Meaning, and Blame*. (Cambridge:

16 It is important to note that Talbert’s main focus is on warranted resentment, and that he, Scanlon and Smith, do not necessarily accept that resentment entails all of the aspects of accountability that are mentioned in Watson’s characterization.


18 The objection just discussed concerned a possible asymmetry between moral virtue and vice. Several people to whom I posed the symmetry argument pointed out that other sorts of asymmetries have been thought to exist in the realm of moral responsibility and that they, too, might be thought to pose additional challenges for the symmetry argument. It has been argued that there is an asymmetry between blameworthiness and praiseworthiness when it comes to the ability to do otherwise; blameworthiness requires such an ability while praiseworthiness does not. (See, for example, Susan Wolf, *Freedom Within Reason*. (Oxford: Oxford University Press, 1990) and Dana Kay Nelkin, *Making Sense of Freedom and Responsibility* (Oxford: Oxford University Press, 2011). However, it is important to note that this particular asymmetry is an asymmetry with respect to accountability, and is motivated in part by its being entailed by a
unified condition for accountability that points to a particular ability to respond to reasons.

Without an independent reason for thinking that an asymmetry is plausible in the case of attributability on the dimension of moral understanding, I do not think a different asymmetry, if there is one, in the case of accountability gives us any reason to do so.

Another kind of asymmetry in the realm of responsibility that has been noted is a so-called “side effect asymmetry.” (See Joshua Knobe, “Intentional and Side Effects in Ordinary Language,” *Analysis* 63 (2003): 190-94.) Subjects judge agents more blameworthy for the unintended, but foreseen, harmful side effects of their actions about which the agents do not care, and much less praiseworthy for the unintended, but foreseen, beneficial effects of their actions, about which the agents do not care. This is another kind of asymmetry between praiseworthiness and blameworthiness, and one might argue that its existence makes more plausible an asymmetry with respect to kindness and cruelty cases. However, the side-effect asymmetry has a natural explanation in terms of the nature of moral principles that is not available to explain an asymmetry between kindness and cruelty. We have a “negative” but perfect duty to avoid causing harm as the result of our actions in certain circumstances (e.g., where the agent has a right not to be harmed), so the agent who foresees harmful side effects of his agency may be understood by the subjects to have violated a duty and for this reason is responsible for the consequences. On the other hand, because an agent who foresees but does not intend or care about good consequences neither fulfills nor “goes beyond” a duty with respect to the consequences, such an agent is not responsible for them. As before, then, the side effect asymmetry appears to relate to accountability, and also has a plausible rationale. Neither of these features appears to be present in the case at hand.
One might ask about any particular trait or manifestation of it, such as contempt, whether it is moral. (Thanks to a referee who raised the question whether contempt is a moral attitude.)

When it comes to contempt, in particular, I think that the answer depends on how we understand the attitude. On at least one understanding, contempt seems to fit well under the moral rubric. Following Mason, contempt can be understood as an attitude “in response to a failing presumed to lessen another’s worth as a person, in the sense of lowering their standing in the system of expectations, demands, and rights (merited and owed) that define normative relations with our fellows…In short…presenting its object as low in the sense of ranking low in worth as a person in virtue of falling short of some legitimate interpersonal ideal of the person, one the contemner endorses…” (Michelle Mason, “Contempt as a Moral Attitude,” *Ethics* 113 (2003): 234-272); (240-41)). On this conception of contempt, contempt seems to require a great deal of moral understanding having to do with one’s value as a person, and to count as a moral attitude. While Mason argues that contempt as an attitude can sometimes be justified, this leaves open that contemptuousness as a character trait or contemptuousness as a disposition to unjustified contempt is a moral vice. Of course, one might use “contempt” to refer to something that requires less in the way of cognitive sophistication, perhaps as something closer in kind to hatred or disgust. In the latter sense, perhaps contempt is not a moral attitude. As long as we are clear about which sense of the term we are using, we can distinguish clearly between the moral and non-moral.

I have taken it to be sufficient for at least some sort of moral understanding that one take others’ interests to be reason-giving. But this is not the only sort of moral understanding. On the characterizations of the psychopaths set out earlier, psychopaths are described as “morally blind” along a number of dimensions, including: not understanding the difference between moral and
conventional requirements and thereby having an understanding that others’ interests are reason-giving in themselves (Talbert), or, similarly, not having an understanding of others’ interests as making valid claims (Watson), not having a conception of justice and the good and not being able to engage in mutual recognition that lie at the core of moral life (Watson). Thus, the moral blindness of the psychopath, at least as originally described, encompasses all of these (overlapping) deficits. Thus, even if we have a minimal threshold for moral understanding, it seems that the psychopath, at least as originally described, will not meet it, and so will not qualify for the moral virtues and vices. Of more general significance, and setting aside the stipulated characterization of the psychopath, the symmetry argument concludes that possession of moral vices and virtues (such as cruelty and kindness in particular) requires at least minimal moral understanding.

Levy (“Responsibility of the Psychopath,” 135) also questions whether psychopaths can exhibit contempt, given that they lack moral knowledge. He argues for the conclusion on the basis of an analogy. Suppose that plants can be harmed and that this provides us with a moral reason against treading on them. Since we are ignorant of this harm, then our treading on plants does not manifest contempt. Similarly, psychopaths’ actions do not express contempt. For a similar argument, see Shoemaker (“Attributability, Answerability, and Accountability,” 628), who interestingly accepts the conclusion that psychopaths do not exhibit contempt, but rejects the idea that they are not responsible in the attributability sense, because he thinks that they still exhibit callousness and cruelty. While I am sympathetic to the shared conclusion here, regarding contempt, the example may need to be modified since as Talbert (“Accountability, Aliens”) points out, there is a disanalogy between the two sorts of cases in that one sort of ignorance is of
consequences (in the plants case) and one is moral, and it is at least an open question without further argument whether this is relevant.

The ultimate plausibility of the reasoning on this horn of the dilemma might very well depend in part on a full adjudication of a debate about the nature of pain (and pleasure). On so-called “evaluative” theories of pain, for example, pain is partly constituted by an evaluative judgment to the effect that there is a harm to the body (Norton Nelkin, “Reconsidering Pain,” Philosophical Psychology 7 (1994): 325–343), or by a felt evaluation to the effect that something bad is happening, which itself entails certain motivational states, such as stopping the cause of it (Bennett Helm, “Felt Evaluations: A Theory of Pleasure and Pain,” American Philosophical Quarterly, 39 (2002): 13–30). A related view, the so-called “imperative” view, take pain to be composed both of judgments of damage to the body and commands such as “stop!” (see, for example, Richard J. Hall, “Are Pains Necessarily Unpleasant?” Philosophy and Phenomenological Research, 49 (1989): 643–659.). For a helpful survey of these and other views about the nature of pain, see Murat Aydede, “Pain”, The Stanford Encyclopedia of Philosophy (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/pain/>. If any in this family of views about the nature of pain are plausible, they could form part of an argument for the conclusion that understanding others’ pain requires understanding of others’ interests as reason-giving. But as a referee rightly points out, it is a further move from necessarily seeing one’s own pain as a reason for one to stop it to seeing someone else’s pain as a reason for one to stop it. As another referee points out, pain asymbolia (a condition of patients who report feeling pain while claiming that it is not unpleasant) would present evidence on the other side. Interestingly, if such patients really do experience and understand pain without believing particular instances of it to be unpleasant,
then while it is true that the second horn of the dilemma is undermined, the first appears to be strengthened. For it is even easier to understand that enjoying the experience of seeing others in pain is not necessarily cruel. While the issues are subtle and interesting, my aim here is not to defend the conclusion of this horn of the dilemma, but to acknowledge the possibility that it is supported by good reasoning, and go on to show that, even if it were correct, it would not help in a defense of premise (1).

Independently of any particular conception of the nature of pain, one might also question the reasoning behind this horn of the dilemma on other grounds, namely, that this understanding of pain, together with an understanding of cruelty as associated with the taking of pleasure in others’ pain, would require too much to achieve cruelty. In particular, it would require that cruel persons take the reasons not to cause pain to be overridden, or at least that they be indifferent to such reasons. (I thank a referee for this point.) My response is that, intuitively, this is not obviously too high a standard for cruelty. In fact, there is some reason to think that we ought to withhold judgment on what exactly cruelty requires until we understand what pain itself requires, or, at least, that we consider these questions together. In that case, if the reasons for the evaluative theories of pain are strong, then perhaps cruelty will simply turn out to require more than it might have seemed.

23 For a subtle and sustained development of the idea of moral accountability as a conversation, see Michael McKenna, Conversation and Responsibility (Oxford: Oxford University Press, 2012).

24 For an illuminating discussion of the distinction between different kinds of reasons one might have for performing or omitting to perform a given action, such as not stepping on another’s foot, see Steven Darwall, “Moral Obligation and Accountability,” Oxford Studies in Metaethics
II (2007): 11-132 and Steven Darwall, *The Second Person Standpoint: Morality, Respect, and Accountability* (Cambridge: Harvard University Press, 2006). One kind is simply a matter of seeing that the other person’s pain is a bad thing while the second is essentially related to the fact that the other person has the authority to demand that you not do so (“Moral Obligation,” 112-113). It might be argued that one could have the capacity to understand the first kind of reason without the capacity to understand the second, and on these grounds that (v) is plausible. Rather than offering a full assessment of this argument, however, I pursue a different approach in the text. At the same time, in defending a different argument, I respond to some objections that appear to apply equally to both.

25 At this point, one might note that the Argument from Intelligibility includes an alternative explanation for why the relevant demands can have reasons-free content but still count as a way of holding morally accountable. The reason is that the moral quality of the demand comes from the recognizable moral authority of the one making the demand, rather than from the propositional content of the demand itself.

26 A second way to see that the demands are demands to act for certain reasons begins with the idea that demands associated with moral accountability are by their nature demands to comply with moral obligations. And moral obligations in turn make reference to agents’ reasons. (See for a defense of this latter claim.) In my view, this line of reasoning ultimately offers a deeper explanation of the connection between accountability and agents’ reasons. But I do not believe it is necessary to accept this explanation in order to accept the conclusion that the demands are demands (in part) to recognize certain reasons.

27 Consistent with the idea that one is not thereby “off the hook” are a variety of different ways of identifying what one is (directly) accountable for. (E.g., is one directly accountable only for

28 See also Smith’s reply in “A Unified Account” to Shoemaker’s “Attributability, Answerability, Accountability” in which she defends a unified account against what Shoemaker sees as three distinct notions of responsibility. A full discussion of Shoemaker’s account must await another occasion.

29 Watson sometimes suggests that the relevant demands are of the sort Smith suggests. For example, recall that he writes, “[B]ecause they are incapable of the reciprocity that demanding and owing justification presumes, moral criticism is not only futile but senseless” (“Trouble,” 314). But here I depart from this reasoning. As I have suggested, the more fundamental—or at least an equally fundamental—reason that psychopaths are not accountable is that they cannot meet the demands that apply at the time of action. That is, they cannot see the reasons on which they ought to act.

30 The Ought Implies Can principle is, of course, controversial. But note that the principle needed here can be weaker than a traditional understanding of the Ought Implies Can Principle on a number of dimensions. First, the principle needed here is in the first instance a principle
about whether it makes sense to make demands if the one to whom the demands are directed lacks the capacity to meet them (rather than a principle directly about obligations or oughts).

Second, although the principle has often been taken to imply the incompatibilism of obligations (or at least those one does not fulfill) and determinism, there are many accounts of the principle available that do not have this implication. (To see this, note that there have recently been a number of compatibilist accounts of the relevant “ability.” See Michael Fara, “Masked Abilities and Compatibilism,” *Mind* 117 (2008): 843-65, Nelkin, *Making Sense*, Michael Smith “Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion,” in Sarah Stroud and Christine Tappolet (eds), *Weakness of Will and Practical Irrationality*, 17-38. (Oxford: Oxford University Press, 2003), Kadri Vihvelin, “Free Will Demystified: A Dispositionalist Account,” *Philosophical Topics* 32 (2004): 427-450, and Wolf, *Freedom Within Reason* for some examples.)

31 Given Talbert’s focus on resentment, and lack of endorsement of the other aspects of accountability described earlier, there is reason to believe he would endorse this reading of the objection.

32 Interestingly, Fischer and Tognazzini argue that attributability itself should be divided into two “stages” or “steps” on the way to “accountability of the strongest sort” (John Martin Fischer and Neil Tognazzini, “The Physiognomy of Responsibility,” *Philosophy and Phenomenological Research* 82 (2011): 381-417.) They distinguish between an aretaic notion that allows for the attribution of excellences and faults, and a reactive notion that allows for the target to be the sensible object of reactive attitudes. They argue that psychopaths might illustrate this distinction in that they are proper targets of aretaic judgments, but perhaps not reactive attitudes. I have argued that—at least as stipulated—psychopaths are not proper objects of appraisal in terms of
moral excellences and faults, and thus, that they would not satisfy even the conditions for the first sort of attributability. Consideration of the hypothetical incorrigible racists, however, suggests that the moral attitudes such as resentment might be apt as responses to moral faults such as cruelty, even if it does not make sense to make certain demands of those with such faults. This is one way in which the significance of the moral/non-moral distinction discussed in section 4 might be manifested.

33 This conclusion is consistent with the claim that both are fully legally responsible. For the capacities to understand the laws and consequences for violating them and the capacities to follow them are quite different from the moral capacities required to act or refrain for particular reasons. And it might be that the former sorts of capacities are possessed by psychopaths and incorrigible racists alike, and also that those capacities are the ones required for legal responsibility. (On this point, see Peter Litton, “Psychopathy and Responsibility Theory,” *Philosophy Compass* 5/8 (2010): 676-88, and David Shoemaker, “Psychopathy, Responsibility, and the Moral/Conventional Distinction,” *The Southern Journal of Philosophy* 49 (2011): 99-124; for an argument that the capacities for moral responsibility are also required for legal responsibility, see Cordelia Fine and Jeanette Kennett, “Mental Impairment, Moral Understanding and Criminal Responsibility: Psychopathy and the Purposes of Punishment,” *International Journal of Law and Psychiatry* 27 (2004): 425-43.

Relatedly, the criterion given to subjects in the study for moral wrongfulness does not capture the kind of reasons-giving knowledge that is of moral significance. While it is defined in opposition to “rules, customs, or laws,” it is still defined in terms of how society would regard a particular kind of act. In other words, knowing what the rules are is a kind of sociological knowledge, whereas having moral knowledge is a kind of normative knowledge, or critical competence. (Thanks to a reviewer for this way of putting the point.) For an interesting discussion of related concerns about the Moral/Conventional test, see Shoemaker “Psychopathy, Responsibility.”