Attributionist Theories of Moral Responsibility
Matthew Talbert

in *The Oxford Handbook of Moral Responsibility* (2022),
edited by Dana Nelkin and Derk Pereboom, pp. 53-70

1. An Austere Approach to Responsibility

Several related factors make it difficult to give a confident account of *attributionism* and the *attributionist* approach to moral responsibility. First of all, the view is a fairly recent innovation (though its central components are not without precedent). It wasn't until 2005[1] that several related perspectives on moral responsibility were collected under the title "attributionism," and this was largely in reference to work that had appeared not much earlier. One might worry, then, that the boundaries of the view are not yet clearly defined, and perhaps that authoritative presentations have yet to appear. Secondly, "attributionism" is used to refer to the views of several authors, but depending on who is using the term, the members of this group are not always the same; moreover, none of these authors defend identical theories, so, again, there's room for disagreement about what the approach is supposed to rule in and rule out. Finally, the so-called "attributionists" have generally been reluctant to accept this label. There's some justification for this tendency since, as I explain below, the name is potentially misleading and has perhaps led to some confusion in the literature.

Despite these concerns, progress can certainly be made in defining the topic of this chapter. I will take "attributionism" to refer centrally to an approach to moral responsibility that comprises the common elements in the perspectives defended by Pamela Hieronymi, T. M. Scanlon, Angela Smith, and myself.[2] The approach can be glossed this way: *assessments of moral responsibility are, and ought to be, centrally concerned with the morally significant features of an agent's orientation toward others that are attributable to her, and an agent is eligible for moral praise or blame solely on the basis of these attributions*.[3] In this context, attributionists often focus on an agent's *evaluative judgments* (or perhaps her "quality of will"), as revealed through her actions, omissions, beliefs, and attitudes. We are interested in these

---

[1] Levy (2005).

[2] Other views described as versions of attributionism include Adams (1985), Arpaly (2003), and Sher (2009). For discussion of the relation between Sher's view and attributionism, see Talbert (2016, 53-56, 148-149).

[3] This is compatible with holding agents morally responsible for morally neutral acts that are attributable to them in such a way that they would have been open to praise or blame if the acts had not been neutral (for a similar point, see Hieronymi 2008, 363 note 13). Throughout most of this chapter I will focus on blameworthiness.

judgments because we are interested in how other agents stand with respect to us and to one another: do their judgments, and the actions they inform, indicate an appropriate degree of respect and concern for us and for others, or not?

Representative summaries of "attributionism" include the following:

> On *attributionist* accounts, an agent is responsible for an action just in case that action is appropriately *reflective* of who she most deeply is.… that is *sufficient* for us to hold her responsible. (Levy 2007, 132; third emphasis added)

> [The view] often labeled 'attributionism,' … asserts that agents are responsible for all of the actions, beliefs, and attitudes, *conscious or not*, that reflect their judgments about what they have reason to do, believe, or feel. (Sher 2009, 120; emphasis added)

> [Attributionists] argue that responsibility for an action or omission … *requires only* that the action be appropriately attributable to the agent …. An action is … attributable to the agent if it is expressive of her attitudes and values. (Levy & McKenna 2009, 115; emphasis added)

Angela Smith notes that what is distinctive about attributionism is *not* its interest in specifying grounds for attributing attitudes, actions, and omissions to agents—many other approaches share this feature (Smith, unpublished).[4] Instead, and as I meant the added emphases above to bring out, the distinctive feature of the view is that the relevant attributions are taken to be *all* that is required for responsibility. The above quotations point, then, to the relative *austerity* of attributionism, to borrow a term that Gary Watson once used to describe Scanlon's approach (Watson 2002, 240).

Attributionism is "austere" in the sense of lacking adornment: it gets by without positing much in the way of necessary conditions on moral responsibility.[5] For example, attributionists tend to reject both the claim that you are morally responsible *only* for what is under your voluntary control, as well as the claim that you are responsible for wrongdoing *only if* you could have responded to the moral considerations that spoke against your behavior.[6] These conditions are

---

[4] I thank Smith for permission to cite and quote her unpublished paper. Some of the themes in this paper find their way into Smith (2015).

[5] In calling Scanlon's view "austere," Watson also meant to suggest that it can seem coolly detached, concerned merely with assessing agents' qualities. I take up this criticism below.

[6] Associated with the rejection of these conditions is a tendency among attributionists to also reject historical conditions on responsibility (i.e., conditions that exempt from blame agents with certain sorts of personal histories) as well as certain requirements on moral knowledge (such as the requirement that blameworthy wrongdoers recognize, or could have recognized, the moral status of their behavior).

rejected for the same reason: agents may reveal features of their selves, in a way that is sufficient for moral responsibility, even in cases in which they are not met.

This suggests another sense in which attributionism can seem "austere." Perhaps it is too stern and severe a view, one that, in rejecting the above conditions, holds too many people responsible. At any rate, this has been one of the central criticisms of the view.[7]

2. The Recent History of "Attributionism"

In her 1990 book, *Freedom Within Reason*, Susan Wolf criticizes what she calls the "Real Self View" of moral responsibility. According to this view, "an agent is responsible only for those actions which are *attributable* to her real self," and whether one's actions are attributable to one's real self depends on whether "one is at liberty to govern one's actions on the basis of one's valuational system" (Wolf 1990, 34, emphasis added). These quotations suggest a connection between the Real Self View and attributionism, but the suggestion is a bit misleading. Certainly, both approaches elucidate the conditions under which behavior is attributable to agents (though, as noted above, this does little to distinguish these views from many competitors), but the Real Self accounts with which Wolf is concerned (e.g., Watson 2004 [1975]) focus on a fairly narrow set of issues having to do with an agent's relationship to her own desires: does she endorse these desires, do they align with her values, or is she alienated from them in a way that calls her responsibility into question? Contemporary attributionists give little detailed attention to this particular issue, focusing instead on developing broader accounts of responsibility and blameworthiness.

However, one of Wolf's criticisms of the Real Self View brings out an important connection with attributionism. For Wolf, a traumatic upbringing can give us reason to "question an agent's responsibility *for* her real self," which may raise doubts about the agent's responsibility for her present behavior (1990, 37). But for the Real Self theorist, as long as an agent has a properly integrated "real self" (as long as she is moved by desires that accord with her values), she will not be exempted from responsibility on the basis of historical considerations: it simply "does not

---

[7] Attributionism is thus criticized for setting the bar for blameworthiness too low, but some versions of the view may also seem to set the bar too high. Suppose that the attribution of a morally faulty attitude is not only sufficient for blameworthiness, but *necessary* as well. The problem is that certain kinds of unwitting wrongdoers (different from the morally ignorant wrongdoers alluded to in the previous note) do not harbor the attitudes that this version of attributionism requires for blameworthiness. For development of the view along these lines see Talbert (2017).

matter where her real self comes from" (Wolf 1990, 35). As we shall see, attributionism reaches a similar conclusion: the fact that one's bad behavior is the predictable result of one's upbringing does not necessarily get one off the hook.

In order to make sense of the discrepancy between her own view and the Real Self theorist's, Wolf suggests that the latter may be working with a relatively superficial conception of responsibility (1990, 39-40). The picture that Wolf paints is of an arid and detached—one might say *austere*—take on moral responsibility: on this view, to blame someone is simply to see her as a producer of bad outcomes, as merely "a bad act-maker" (1990, 39). But, Wolf argues, simply identifying an action as bad, and an agent as its author, misses the depth and significance of our moral responsibility practices (1990, 40-41).

In an important 1996 paper, Gary Watson responds to Wolf. He agrees that some perspectives on moral responsibility focus on the nature of agents' actions and on the attributability of these actions to their respective agents. Watson calls these "self-disclosure" views: when an agent acts in a way that he reflectively endorses, this discloses something important—and relevant to responsibility—about his "fundamental evaluative orientation" (2004 [1996], 271). Watson contends, against Wolf, that this is not a superficial form of moral responsibility: it is, rather, "a core notion of responsibility that is central to ethical life and ethical appraisal" (2004 [1996], 263). But Watson concedes to Wolf that identifying instances of self-disclosure, and the moral quality of these disclosures, is not *all* that there is to moral responsibility:

> Holding people responsible is not just a matter of the relation of an individual to her behavior; it also involves a social setting in which we demand (require) certain conduct from one another and respond adversely to one another's failures to comply with these demands. (2004 [1996], 262)

For Watson, the self-disclosure theorist's focus on attributability is associated with a lack of attention to the conditions an agent must fulfill in order to be properly *held responsible* for his behavior. For example, self-disclosure views "are silent … about the capacity of self-governing agents to comprehend the grounds on which moral requirements rest, and about our authority to hold one another to these" (2004 [1996], 262-63). Thus, following Wolf, Watson suggests that self-disclosure views lose sight of the fact that agents formed by unfortunate circumstances might not be properly held responsible for bad behavior that is admittedly attributable to them. These agents may act badly, but without being blameworthy in a sense that goes beyond attribution "to

involve the idea that agents deserve adverse treatment or 'negative attitudes' in response to their faulty conduct" (2004 [1996], 266).

Watson is distinguishing here, as he puts it in the title of his paper, between *two faces of responsibility*. A person is morally responsible in the *attributability sense* if she satisfies conditions that allow us to attribute behavior to her. But this person may fail to satisfy the more demanding conditions—historical conditions, moral competence conditions, etc.—that apply to responsibility in the *accountability sense*. Such an agent may be open to aretaic appraisal (virtues and vices may be attributed to her), but not properly held accountable for her behavior and targeted with the negative responses that seem to be at the heart of blame.

What does this have to do with contemporary attributionism? One worry is that readers may conflate the contemporary view with Watson's "*responsibility as attributability*" (2004 [1996], 271). This conflation is understandable: both views take moral responsibility to depend, roughly, on morally significant disclosures of an agent's *fundamental evaluative orientation* (as we saw Watson put it) or of her "moral personality" (Hieronymi 2008). In addition, T. M. Scanlon, for one, does refer to the sort of responsibility that is relevant to moral blame as "*responsibility as attributability*" (1998, 248)*,* and he tends to speak interchangeably of "moral criticism" and "blame," which suggests a deflated account of blame to many readers (1998, 268). However, Scanlon also says, "when moral criticism applies this makes various reactive attitudes such as guilt, resentment, and indignation appropriate" (1998, 276). It is not correct, then, to suppose that Scanlon is concerned just with specifying the conditions for *attributing* moral faults to agents, and not with drawing conclusions about whether agents are open to the responses involved in holding agent's accountable for their behavior. The same goes for the other attributionists considered in this chapter: they take themselves to be specifying grounds for holding agents accountable for their behavior, it's just that these grounds are found in the fulfillment of attributability conditions. In other words, *attributability is enough for accountability*.[8]

To repeat, the worry about conflation is, as Angela Smith puts it, that "when Scanlon says he is presenting an account of responsibility in the 'attributability' sense, one will mistakenly think

---

[8] Michael McKenna says that Angela Smith "collapses Watson's distinction by, so to speak, downgrading all cases of accountability-responsibility to attributability-responsibility" (2012, 193). It would be better to say that attributability-responsibility is *upgraded*. I take Pamela Hieronymi to be making a related point when, in her discussion of Watson's distinction, she says, "I will suggest that the reactive attitudes cannot be separated from 'aretaic' appraisal" (2004, 140 note 12).

that he is offering an account of the conditions of [just] aretaic appraisal" (Smith, unpublished). And Smith suggests that we have an example of this mistake in the paper that first described Scanlon's view, and Smith's own, as instances of "attributionism": Neil Levy's "The Good, The Bad and the Blameworthy." There, Levy says that Watson's distinction between *attributability* and *accountability* "seems to map neatly onto the distinction between responsibility as understood by attributionism and responsibility as understood by volitionism" (2005, 3).[9] Smith worries that this "falsely implies that … [views like hers] are concerned only with the conditions for … aretaic appraisal," whereas Smith aims to provide conditions for more substantial blaming responses (Smith, unpublished).

Smith argues that this worry—together with her point, mentioned above, that attributionists are not alone in their interest in the attribution of behavior—should lead us to dispense with the label "attributionism." Smith is right that it is an error to run together contemporary attributionism and the view that Watson described. However, for several reasons, I've come to think that we should be content with "attributionism" as the name for the view discussed in this chapter.

For one thing, I don't think that the error in question has significantly derailed debate. Debate would be derailed if writers failed to engage attributionism because they mistakenly believe that the view is not concerned with giving an account of *real* (i.e., resentment-involving) moral blame. But this isn't what has happened. Even in the paper to which Smith refers, Levy argues that attributionism generates the wrong result—one that affirms responsibility—in cases of impaired agents. As I read him, Levy means that attributionism regards these agents as open to something like accountability-blame when they should not be so regarded. He argues, for instance, that, contrary to the attributionist account, a suitably impaired agent may plausibly be "taken to be a disagreeable character, [but] he is not therefore a blameworthy person" (2005, 8). I take Levy to be saying that negative aretaic appraisals of such an agent may be appropriate, but he will not be blameworthy in some other, more robust way. But this makes sense *as a reply to attributionism* only if Levy takes that view to hold that impaired agents may be accountable in a way that goes beyond aretaic appraisal. Generally, this is what we find in the literature. As I noted in the last section, critics regard attributionism as inclined to hold *too many* people responsible. But this isn't supposed to be a problem because the view engages too readily in mere aretaic appraisal; it's supposed to be a problem because the view blames too many people in a more profound way.

---

[9] "Volitionism" is Levy's term for views that place a voluntary-control condition on responsibility.

As long as those who argue against attributionism are arguing about the appropriate grounds for serious moral blameworthiness, then the mistake described above hasn't seriously distorted the debate. A couple other considerations speak in favor of preserving "attributionism." First, this ship seems to have sailed: the term is fairly well ensconced in the literature. Second, alternative proposals have their own problems. Smith sometimes substitutes "non-volitionalism" for the view that others call "attributionism" (2008), but this obscures the degree to which the view rejects conditions on responsibility besides those having to do with volitional control. And while Levy has substituted talk of "quality of will" views for "attributionism" (2011, 158), too many other accounts—ones that are clearly not varieties of attributionism—can lay equal claim to this title: for example, McKenna (2012), Shoemaker (2015), and Wallace (1996). Finally, Smith has recently turned to speaking of "answerability" because, on her account, people are morally responsible for that for which they can reasonably be asked to answer (2015; see also Hieronymi 2008 and 2014). However, "answerability" suffers from a version of the problem that "attributionism" has: some authors (Shoemaker 2011 and 2015 and Pereboom 2014) use the term to describe a type of responsibility associated with non-resentment-involving blame. And, as Smith herself notes, it's sometimes simply easier "to make certain points using the language of 'attributability'" (2017, 37 note 2). I propose, then, that we so-called "attributionists" accept this title.

3. The Voluntary Control Condition

It's often assumed that we are blameworthy only for what is in our control, either in an immediate or a mediated fashion. After all, many of the things for which we *are* potentially blameworthy—our conscious choices, for example—have this feature. However, attributionism holds that we are open to blame on account of those things that reflect our objectionable evaluative judgments and that not all such things are under our immediate control or are associated with prior exercises of control.

In *What We Owe to Each Other*, Scanlon argues that blameworthiness tracks instances of morally faulty self-governance:

> If an action is blameworthy …. the agent's mode of self-governance has ignored or flouted requirements flowing from another person's standing as someone to whom justification is owed. This is what … makes it appropriate for the person who was wronged to feel resentment rather than merely anger and dismay. (1998, 271)

Understood in this way, blame—though, again, Scanlon often speaks simply of "moral criticism"—will be appropriate only with respect to agents who are capable of governing themselves in the face of reasons. And in the case of these agents, such a response will be apt "only in regard to their judgment-sensitive attitudes: that is, those attitudes that … should be 'under the control of reason'" (Scanlon 1998, 272).[10]

Of course, not all criticism is *moral* criticism since not all reproaches call into question a person's judgment sensitive attitudes. There might be a criticism implicit in a question like "'Why are you so tall?,'" but it's not a moral criticism—at least not an apt one—since a person's height does not implicate morally significant judgments on her part (Scanlon 1998, 272). It's also true, of course, that a person does not choose her height. But we should not conclude from this that we reasonably blame people only for what they choose. In fact, some features of our selves that we do not choose—and that we might not be able to change—may reflect our judgments, and so we may be open to blame on their account. Thus, on Scanlon's view, it is not the case "that moral criticism applies only to actions or attitudes that arise directly from an agent's conscious judgments" (1998, 272; also see Scanlon 2008, 170). Such criticism can also apply to involuntary attitudes that are judgment-sensitive.

Angela Smith has developed this aspect of attributionism with particular force (also see Hieronymi 2008 and 2015), emphasizing the ways in which unchosen features of our selves—attitudes, desires, what we notice or fail to notice—can reflect our commitments. We may not be causally responsible for these facts about our selves, but we are responsible for them in the sense that they reflect how we are oriented toward others, and we may properly be asked to alter or defend this orientation.

What Smith is noticing is that there is a *rational relation*[11] between various involuntary states of agents and what they take to be important or valuable. There is,

> a rational connection between many of the thoughts and desires that occur to us and
> the evaluative judgments and commitments we accept. If we value something and
> judge it to be worth promoting, protecting, or honoring in some way, this should

---

[10] Don't be misled by Scanlon's reference to "control." Attitudes are under the control of reason if they are sensitive to judgments about reasons, but these judgments themselves need not be the result of a volitional process (see Scanlon 1998, 281). See note 12 for a related point.

[11] Smith has referred to her own account, therefore, as the "rational relations view" (2005, 240).

(rationally) have an influence on our unreflective patterns of thought and feeling. (2005, 247)[12]

For the attributionist, the presence of this relation explains why we hold one another responsible for what we do voluntarily, and since the same relationship can hold in the case of things that are not voluntary, we can be blameworthy for some things that are not voluntary. If a person's (involuntary) indignation indicates an objectionably contemptuous judgment, or her (involuntary) inadvertence indicates an objectionable lack of concern, then she is open to moral criticism (blame), which responds just to the "content of that attitude and not to facts about its origin in a person's prior voluntary choices, or to facts about its susceptibility to influence through a person's future voluntary choices" (Smith 2005, 251).

Scanlon makes related points in the context of discussing manipulation scenarios of the sort that incline some to put historical conditions on moral responsibility. Of course, some varieties of manipulation do undermine responsibility. If a person's behavior is brought about by hypnosis or direct brain stimulation, then it is unlikely that he is morally responsible for that behavior (Scanlon 1998, 277). However, Scanlon argues that this isn't because the manipulated agent was subject to causal pressures over which he lacked control. Rather, responsibility is absent in these cases because the causal pressures at issue "are of a kind that sever the connection between the action or attitude and the [manipulated] agent's judgments and character" (Scanlon 1998, 278). If an agent's action is brought about by direct brain stimulation, there are no grounds for "attributing to him the discernment or lack of discernment that would be revealed by [the agent] thinking that he had good reason" to perform the action and doing so on that basis (Scanlon 1998, 278).

But what if an action is brought about "by 'implanting' in the agent the thought that it is warranted" (Scanlon 1998, 278)? Such an agent may still be relieved of responsibility even though the action in question expresses an objectionable judgment, but, once again, this is not mainly because of lack of control on the agent's part. Instead, responsibility is called into question to the degree that the judgment is aberrant and transitory, and thus not fully attributable to the agent.

However, in cases of longer-term involuntary changes, our conception of what belongs to the agent for purposes of moral assessment may shift. A previously kind person might undergo a

---

[12] Smith uses "judgment" loosely: judgments "are not necessarily consciously held propositional beliefs, but rather tendencies to regard certain things as having evaluative significance…. 'Judgments' in this sense do not always arise from conscious choices or decisions…" (2005, 251-252).

change of character over which she lacks control "after being hit on the head or given drugs for some medical condition" (Scanlon 1998, 279). Here, "at least at first," we might not see the agent's out-of-character behavior "as grounds for modifying our opinion of her" (Scanlon 1998, 279). But if the pattern of behavior persists, we may come to see it as a part of the agent *as she is now*, and as forming an appropriate basis for moral criticism going forward (Scanlon 1998, 279; also see Smith 2005, 261 note 46).

Some readers may think that if a person owes her bad character to a bump on the head, or to a bad upbringing, then she cannot be open to serious moral blame. Perhaps it is not enough for blameworthiness that an agent's actions reflect her bad attitudes and judgments; perhaps blameworthy agents must also be (in a causal sense) responsible for these facts about themselves. Neil Levy, for example, holds that "I am responsible for my attitudes if I have genuinely been … active with regard to them; if I have chosen them," and he objects to the attributionist's contention that "[i]t need not be the case that I have controlled [my attitudes], or even that I could control them, for me to be responsible for them" (2005, 10). Levy concludes, "[i]f control matters, then its absence cancels responsibility (unless of course the agent is responsible *for* her absence of control) (2005, 10).[13]

According to Levy, what leads the attributionist astray here is a failure to acknowledge the distinction between agents that are bad and agents that are blameworthy. For the attributionist, if objectionable attitudes are attributable to an agent, then she is blameworthy, but, Levy argues, "even when attitudes are rightly attributed to agents, it is a further question whether they are responsible for them" (2005, 15). And with respect to the lapses of attention and concern that Angela Smith describes, Levy says,

> unless we exercise relevant control over them, we are not responsible for our lapses. Since there seems to be conceptual room for a distinction between a *faulty* attitude … and one for which the agent is *at fault*, and attributionists have given us no good reason for thinking that this distinction should not be made, we should reject attributionism as an account of moral responsibility." (2005, 15; see Smith 2008 for her reply)

Despite Levy's suggestion, attributionists are willing to grant that, beyond establishing the attributability of an attitude to an agent, there is also the question of whether the agent is *responsible for* her attitudes in the sense of having played a certain sort of role in their acquisition.

---

[13] Levy, I take it, affirms the antecedent of this conditional.

But attributionists take the attributability of the attitude to settle—if it is suitably objectionable—the question of the agent's blameworthiness even if she is not responsible for bringing it about that she has that attitude. Likewise, the attributionist is happy to distinguish between a *faulty* attitude and an attitude for which an agent is *at fault*. But, again, the attributionist insists that an agent's being *at fault* for an attitude is not required for it to be faulty, and that the faultiness of the attitude is enough for an agent to be open to blame on account of it (even if it is not her fault—causally speaking—that she has that attitude). Levy and others (e.g., Rosen 2004) appear to believe that being at fault for one's faulty attitudes is required for blameworthiness. From the attributionist's perspective, this runs together questions about causal responsibility and moral responsibility in an unhelpful way.

4. The Moral Competence Condition

Many accounts of responsibility hold that wrongdoers are open to serious moral blame only if they are *morally competent* in such a way that they *could have* recognized, and responded appropriately to, the moral considerations that counted against their behavior. [14] Again, attributionism rejects this condition on responsibility.

In *What We Owe to Each Other*, Scanlon says that a "plausible test" for deciding whether a certain type of impairment of moral competence rules out moral criticism is to think about "whether the behavior of a creature which has that condition would, for that reason, lack the distinctive significance that moral failings generally have for relations with others" (1998, 287-88). So, for example, if a being is non-rational and does not make judgments about reasons at all, then, though it may harm us, its behavior would not have the significance that wrongdoing typically has for us. This is because such a being's behavior would not express a judgment about the sort of treatment to which we are open: its behavior would not be of the sort that can "challenge our moral standing and make resentment an appropriate reaction" (Scanlon 1998, 288). However, Scanlon argues, "a rational creature who fails to see the force of moral reasons," but who is otherwise sensitive to rational considerations, is capable of expressing objectionable evaluative judgments, and is, therefore, open to moral blame (1998, 288).

---

[14] For different versions of this condition, see McKenna (2012), Nelkin (2011), Wallace (1996), and Wolf (1987, 1990), among others.

Some agents have localized impairments of moral competence. In this context, the debate often focuses on individuals who form inaccurate moral judgments because those judgments are widely accepted and nurtured in their cultural contexts. One side in this debate holds, roughly, that widespread acceptance, and cultural sanctioning, of an objectionable judgment can impair moral competence insofar as agents in these contexts often have difficulty assessing the moral status of the judgment in question. This can make it unreasonable to expect an agent to avoid wrongdoing associated with this objectionable judgment.

In a well-known discussion along these lines, Susan Wolf considers "persons who, though acting badly, act in ways that are strongly encouraged by their societies—the slaveowners of the 1850s, the Nazis of the 1930s, and many male chauvinists of our fathers' generation, for example" (1987, 56-57). These individuals had the misfortune of being raised in environments that cultivated bad values and bad judgmental tendencies in them, which made it difficult for them to recognize that some of their behavior was objectionable. Wolf concludes that if we believe that these "agents could not help but be mistaken about [the moral status of] their values"—if we believe that they were unavoidably impaired in this way—then "we do not blame them for the actions those values inspired" (1987, 57).

The attributionist has a number of replies to this line of thought (some of which I will save for the discussion of resentment and protest in the next section). To start with, it should be emphasized that when we consider the racist behavior of a morally benighted slaveowner, we are considering voluntary behavior that is often conducted in full awareness of the disagreeable consequences of that behavior for others: the slaveowner simply dismisses—or at least he does not assign appropriate weight to—these considerations. Thus, it should be admitted by all sides that the actions of such a slaveowner can display objectionable attitudes that are at least very similar to the kinds of attitudes that are often relevant to judgments of blameworthiness.

Those who reject attributionism will maintain, however, that while the expression of such attitudes is often associated with blameworthiness, it is not sufficient for blameworthiness: in addition, blameworthiness requires (for example) that it is reasonable to expect a wrongdoer to have avoided acquiring the attitudes for which we blame her. But now the attributionist might argue that it is not clear how adding the satisfaction of this condition (or related ones) to the story of a slaveowners' wrongful behavior would change that behavior from something that does not warrant blame to something that does. Such an addition would be relevant if the fact that a

slaveowner *might have* come to a different conclusion about, say, the moral status of slavery means that his behavior is morally objectionable in a way that it could not otherwise be. But it is not clear that this is so: it is not clear that the behavior of slaveowner *A* is objectionable in some special way, as compared to the behavior of similarly-motivated slaveowner *B*, simply because *A could have* (in some sense) judged that slavery is wrong. Or, more modestly, even if *A*'s behavior would be made especially objectionable by access to this alternative judgment, it doesn't follow that *B*'s behavior is not sufficiently objectionable to make blame appropriate in his case (Talbert 2012a, 98-101).

But this doesn't get to the heart of the position of someone like Wolf. The central idea isn't that a slaveowner's behavior would be more objectionable if he could have been receptive to a different moral outlook. The thought is, rather, something like this: if a slaveowner's moral competence were unimpaired, then he would be a fair target for blame because it would be reasonable to expect him to have avoided acquiring the bad values that he possesses and that inform his wrongful behavior. Put the other way around, the thought is that it is unfair to blame an agent if it was unreasonably difficult for her to avoid the behavior for which she is blamed.

Attributionists will agree that it is sometimes unfair to blame those who have significant difficulty avoiding wrongdoing. But, the attributionist will note, this is most clearly the case when the difficulty in question is associated with a failure of self-governance. A coerced agent, for example, may be blameless and may have difficulty avoiding wrongdoing, but the blamelessness here stems from the fact that her behavior doesn't express the evaluative commitments that it would have expressed had she been acting as she pleased.[15]

Quite often, an agent who has difficulty avoiding an action fails (like the coerced agent) to govern her behavior as she sees fit, but these two things can come apart: an agent may be unable to avoid wrongdoing, but not because her control over her behavior is impaired. For example, a slaveowner's difficulty avoiding wrongdoing may stem from an entrenched commitment to seeing members of a certain race as open to being treated like property. Such an agent controls her behavior as she sees fit, and her difficulty avoiding wrongdoing is a function of an abundance of the sort of thing (a morally reprehensible point of view) that often grounds blame. The attributionist

---

[15] A coerced agent may "act as she pleases" (or least as she chooses) in one sense, but in another very clear sense, she is not doing what she wants.

concludes that it is much less clear in this type of case (than it is in the case of a coerced agent) that it is unfair to blame one who has significant difficulty avoiding wrongdoing.[16]

So far, I have focused on limited impairments of moral competence, but what about wrongdoers with broader impairments? Discussions of this topic have focused on *psychopathy*— or at least a philosophical approximation of this disorder. For the purposes of this chapter, it is enough to say that a psychopath is unable to make use of moral concepts and judgments (in the normal way) and unable to find motivation in moral considerations (in the normal way). As such, the psychopath is quite different from the slaveowner considered above. The slaveowner may recognize that there is such a thing as "moral standing," so it is straightforward to attribute to him, and to see his behavior as expressing, judgments about moral standing. But the psychopath, as conceived here, is a figure who is not in the business of using moral concepts at all, which makes it difficult to interpret his behavior as expressing judgments about standing.

Neil Levy argues on this basis that the psychopath's behavior should not be thought of as "expressing contempt, ill-will, or moral indifference" of the sort that would, on the attributionist approach, ground blame (2007, 135). "Contempt," Levy maintains, "is a thoroughly moralized attitude; only a moral agent [which the psychopath is not] is capable of it" (2007, 135; also see Shoemaker 2011, 629). The psychopath certainly fails to express appropriate regard for others, but since success in this domain was never an option, his failure of regard is not an expression of *disregard*.

Attributionists, however, have argued that psychopaths' behavior can have blame-grounding moral significance even if they cannot form judgments with explicit moral content. This argument (Talbert 2008, 2012b, 2014) takes its cue from Scanlon's observation that, "[a] person who is unable to see why the fact that his action would injure me should count against it still holds that this *doesn't* count against it" (1998, 288). The psychopath may not make moral judgments, but insofar as he makes judgments about reasons at all, he dismisses factors that ought to inform one's practical judgments. Thus, the psychopath inhabits a perspective according to which the suffering of others may be overlooked. We can, for example, attribute to him the judgment, "The

---

[16] See Scanlon (1998, 288) and Talbert (2013, 237-238). This reasoning is related to Frankfurt's (1998 [1969]) argument that access to behavioral alternatives is not required for moral responsibility.

fact that this will hurt you is no reason to refrain," and the attributionist argues that this judgment is contemptuous in a way that is appropriately met with moral blame.[17]

But what is *the point* of blaming a psychopath? A number of philosophers emphasize the conversational aspects of our blaming practices. Blame seems not only to communicate moral condemnation and moral demands, but to also invite replies from wrongdoers: they should come to share our views, to apologize, reform themselves, and so forth. But if blame has an essential conversational point, then perhaps it is appropriate only in contexts in which such communication is possible, and the psychopath is not a potential partner in moral conversation.

Gary Watson has emphasized the way in which the attitudes involved in blame often play this kind of conversational role (see also McKenna 2012 and Shoemaker 2015). Watson has said that the "negative reactive attitudes express a *moral* demand, a demand for reasonable regard," and that "[t]o be intelligible, demanding presumes understanding on the part of the object of the demand" (2004 [1987], 230). In a recent discussion of psychopathy, Watson adds that resentment, in particular, commits us "to the appropriateness of [taking up] an inherently communicative stance" toward the target of these responses (Watson 2011, 328 n. 35). The conclusion to draw, Watson suggests, is that since we cannot hope to communicate our moral perspective to the psychopath, we cannot regard him as an appropriate target of resentment.

Attributionism has a response here, one which will become apparent in the next section in which I discuss the roles that protest and resentment play in the attributionist perspective.

5. Protest, Resentment, and Relationships

Attributionists are united in the view that blame is not, in itself, a sanction or a punishment. As Scanlon puts it, "moral blame is fundamentally a judgment of condemnation, not a penalty" (1998, 267). Though it is often unpleasant to be the object of moral criticism and blame, that is not their purpose, nor are they tactics deployed "in order to enforce norms of behavior" (Scanlon 1998, 285). Thus, judgments of blameworthiness are measured by different standards, and answer to different constraints, than decisions to impose hardships on others.

---

[17] Angela Smith (2015, 118) appears to endorse this conclusion about psychopaths, and Pamela Hieronymi suggests, more generally, that blame is not rendered unfair by the fact that an agent is not "capable of controlling her behavior by the light of moral reasons" (2004, 126).

A similar contrast can be drawn between judgments of blameworthiness and expressions of these judgments. And once we make this distinction, concerns about the fairness of blame may seem less pressing since, as Angela Smith argues, they may have more to do with the propriety of expressing our judgments of blameworthiness than with the accuracy of these judgments. The problem, Smith suggests, is that responsibility theorists often mistakenly suppose that if it is wrong "to express moral criticism or to sanction people for their desires, emotions, and other attitudes," then "they cannot *be* morally responsible for them" (2008, 379; also see Smith 2007).

Making a related point, Pamela Hieronymi notes that the negative attitudes involved in blame "are not action-like. That is, they are not voluntary responses to a judgment in the way an intentional action might be" (2004, 120). So, again, whether one ought to hold these attitudes will not be measured by the same standards as *a decision* to harm a wrongdoer. Indeed, as Hieronymi argues, "adopting the reactive attitudes could be rendered unfair only by considerations that bear on the content of the judgments they reveal" (2004, 133). And the judgment that is revealed by an attitude like resentment is something along the lines of "I was wronged" or "I have been disrespected." Various considerations might show that such judgments are erroneous, but the considerations at issue in debates about the fairness of blame are not typically of this sort.

On Hieronymi's account, resentment's function is to mark the fact that one was wronged. In contrast with Watson's conversational approach, nothing about this picture of resentment commits one to viewing the object of resentment as a potential moral interlocutor. Our private resentment need not commit us to the possibility of moral understanding on the part of the one we blame, and even expressions of resentment need not take such an end to be achievable. Of course, the attributionist does not deny that we often hope that our expressions of resentment will inspire contrition, apology, and reform in a wrongdoer. The attributionist's point is only that resentment— or at least a closely related, morally serious blaming emotion—can be appropriately experienced in the absence of these possibilities (for discussion of the "closely related" emotion, see Hieronymi 2014, 31-32 and Talbert 2014, 289-291).

But still, what is *the point* of resentment when we cannot inspire moral understanding in a wrongdoer? It's tempting to say that, for the attributionist, resentment has no point beyond its sensitivity to the fact that one's standing has been called into question. But a bit more can be said. For example, Hieronymi argues that "resentment is best understood as a *protest*" against a wrong done to you (2001, 546). The fact of your having been wronged makes the claim "that you can be

treated in this way, and that such treatment is acceptable," and "[i]n resenting … you challenge" this claim (Hieronymi 2001, 546). I have also argued that we should think of blame and resentment as a form of protest, and that doing so will help the attributionist respond to those who insist that expressions of blame lose their sense if they cannot be interpreted as communicative efforts (Talbert 2012a). Going further, Angela Smith has identified the "desire to protest and repudiate conduct" as "[t]he *fundamental* motivational element underlying all instances of moral blame" (2013, 36-37; emphasis added).

One of Smith's targets in making this last point is Scanlon's account of blame, which brings me to the final issue I want to discuss: Scanlon's perhaps problematic stance with respect to the harsher side of blame, particularly resentment. If Scanlon's account of blame is one that does not go in for resentment (and similar responses), then this again raises the concern (described in Section 2) that attributionism—or at least Scanlon's version of it—is not defending a serious form of blame.

In *Permissibility, Meaning, and Blame*, Scanlon says that a claim of blameworthiness asserts that an agent's "action shows something … that impairs the relations that others can have with him or her" (2008, 128). Further, "[t]o *blame* a person is," in addition to forming a judgment of blameworthiness, "to take your relationship with him or her to be modified" (Scanlon 2008, 128), and "to hold the attitude toward him or her that this impairment makes appropriate" (Scanlon 2008, 131). Attitudes like *resentment* are not a necessary component of blame on this account. Scanlon imagines a case in which his friend Joe betrays him. In response, Scanlon might "decide not to rely on or confide in Joe … and not to seek his company … or to have … special concern for his feelings and well-being" (2008, 136). Scanlon takes all these ways in which he might revise his relationship with Joe to count as forms of blame. In addition, Scanlon might "also resent [Joe's] behavior," yet "this is not required for blame … I might just feel sad" (2008, 136).[18]

Perhaps Scanlon is right that there is more to blame than resentment, but should we agree that his sadness at being mistreated by Joe is, as he suggests, an instance of blame? Surely this is in tension with much of our ordinary talk about "blame" (Wallace 2011; Wolf 2011). And even if

---

[18] It's worth emphasizing that Scanlon makes room for resentment even in his recent work on blame:

> It would be foolish to deny that moral emotions such as resentment … are appropriate responses … [to blameworthy actions]. But these emotions are not all that blame normally involves. Other attitudes, such as modified intentions, are also important. (2013, 99)

we count Scanlon's decision to not seek out Joe's company as a form of blame, it is not the robust sort that is typically at issue in debates about moral responsibility. After all, even someone who is largely skeptical about blame, as conventionally understood, can accept that it is reasonable for Scanlon to choose to spend less time with Joe (Pereboom 2014, 131-132).

Something seems to be missing from Scanlon's account. But the missing element may not be resentment. Instead, I'm inclined toward Smith's proposal that what is missing is *protest*. Perhaps, Smith says, "[i]t is only those modifications of attitudes that are undertaken as a way of *protesting* the relationship-impairing attitudes of others that qualify as instances of moral blame" (2013, 39). Of course, resentment is a common way of registering moral protest, but it is not the only way. Playing off Scanlon's example, Smith says,

> I may have lost my ability to feel anger toward an unreliable friend, yet I may still protest his treatment of me by cutting off relations with him. In doing so, and doing this *in protest* of his latest let-down, I make clear that I blame him, even if my predominant feeling is one of sadness. (2013, 41)

Conclusion

I have characterized attributionism as austere because of its narrow focus on the quality of a specific range of evaluative judgments that are attributable to agents. Successful attribution of these judgments is enough to make an agent blameworthy, regardless of whether he fulfills a number of other proposed conditions on moral responsibility. However, attributionism may also be thought austere in the sense that it is interested only in detached appraisal of agents, or in the sense that it is excessively severe and heedless of constraints that rightly apply to our blaming practices.

In response to these criticisms, attributionism clarifies its focus on evaluative judgments and the attitudes that are sensitive to these judgments. Since the judgments that attributionism appraises are of paramount interpersonal significance, blaming attitudes are rightly in play as responses to these appraisals. Moreover, the blaming responses the attributionist envisions are not interpreted as sanctions, but rather as natural responses to the quality of the evaluative judgments attributed to agents. The fairness of a reactive response depends, then, on the accuracy of the attribution of an objectionable judgment, and the degree to which the response is attuned to, or calibrated by, the attributed judgment.[19]

---

[19] I thank Randolph Clarke and Angela Smith for their helpful comments on a draft of this chapter.

References

Adams, Robert. 1985. "Involuntary Sins." *Philosophical Review* 94: 3-31.

Arpaly, Nomy. 2003. *Unprincipled Virtue: An Inquiry Into Moral Agency*. New York: Oxford University Press.

Frankfurt, Harry. 1998 [1969]. "Alternate Possibilities and Moral Responsibility." In H. Frankfurt, *The Importance of What We Care About*. Cambridge: Cambridge University Press, pp. 1-10.

Hieronymi, Pamela. 2001. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62: 529-555.

Hieronymi, Pamela. 2004. "The Force and Fairness of Blame." *Philosophical Topics* 18: 115-148.

Hieronymi, Pamela. 2008. "Responsibility for Believing." *Synthese* 161: 357-373.

Hieronymi, Pamela. 2014. "Reflection and Responsibility." *Philosophy and Public Affairs* 42: 3-41.

Levy, Neil. 2005. "The Good, the Bad and the Blameworthy." *Journal of Ethics and Social Philosophy* 1: 1-16.

Levy, Neil. 2007. "The Responsibility of the Psychopath Revisited." *Philosophy, Psychiatry, and Psychology* 14: 129-138.

Levy, Neil. 2011. *Hard Luck*. New York: Oxford University Press.

Levy, Neil and Michael McKenna. 2009. "Recent Work on Free Will and Moral Responsibility." *Philosophy Compass* 4: 96-133.

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

Nelkin, Dana. 2011. *Making Sense of Free Will and Responsibility*. New York: Oxford University Press.

Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.

Rosen, Gideon. 2004. "Skepticism about Moral Responsibility." *Philosophical Perspectives* 18: 295-313.

Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, and Blame*. Cambridge, MA: Harvard University Press.

Scanlon, T. M. 2013. "Interpreting Blame." In D. Coates and N. Tognazzini (eds.), *Blame: Its Nature and Norms*. New York: Oxford University Press, pp. 84-99.

Sher, George. 2009. *Who Knew?: Responsibility Without Awareness*. New York: Oxford University Press.

Shoemaker, David. 2011. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility" *Ethics* 121: 602–632.

Shoemaker, David. 2015. *Responsibility from the Margins*. New York: Oxford University Press.

Smith, Angela. Unpublished. "The Myth of Attributionism."

Smith, Angela. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115: 236-271.

Smith, Angela. 2007. "On Being Responsible and Holding Responsible." *The Journal of Ethics* 11: 465-484.

Smith, Angela. 2008. "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 138: 367-392.

Smith, Angela. 2013. "Moral Blame and Moral Protest." In D. Coates and N. Tognazzini (eds.), *Blame: Its Nature and Norms*. New York: Oxford University Press, pp. 27-48.

Smith, Angela. 2015. "Responsibility as Answerability." *Inquiry* 58: 99-126.

Smith, Angela. 2017. "Unconscious Omissions, Reasonable Expectations, and Responsibility." In D. K. Nelkin and S. C. Rickless (eds.), *The Ethics and Law of Omissions*. Oxford University Press, pp. 36-60.

Talbert, Matthew. 2008. "Blame and Responsiveness to Moral Reasons: Are Psychopaths Blameworthy?" *Pacific Philosophical Quarterly* 89: 516-535.

Talbert, Matthew. 2012a. "Moral Competence, Moral Blame, and Protest." *The Journal of Ethics* 16: 89-109.

Talbert, Matthew. 2012b. "Accountability, Aliens, and Psychopaths: A Reply to Shoemaker," *Ethics* 122: 562-574.

Talbert, Matthew. 2013. "Unwitting Wrongdoers and the Role of Moral Disagreement in

Blame." In D. Shoemaker (ed.), *Oxford Studies in Agency and Responsibility*, Vol. 1. Oxford: Oxford University Press, pp. 225-245.

Talbert, Matthew. 2014. "The Significance of Psychopathic Wrongdoing." In T. Schramme (ed.), *Being Amoral: Psychopathy and Moral Incapacity*. Cambridge MA: MIT Press, pp. 275-300.

Talbert, Matthew. 2016. *Moral Responsibility: An Introduction*. Cambridge, UK: Polity Press.

Talbert, Matthew. 2017. "Omission and Attribution Error." In D. K. Nelkin and S. C. Rickless (eds.), *The Ethics and Law of Omissions*. Oxford: Oxford University Press, pp. 17-35.

Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Wallace, R. Jay. 2011. "Dispassionate Opprobrium: On Blame and the Reactive Sentiments," In R. J. Wallace, R. Kumar, S. Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. New York: Oxford University Press, pp. 348-372.

Watson, Gary. 2002. "Contractualism and the Boundaries of Morality: Remarks on Scanlon's *What We Owe to Each Other*." *Social Theory and Practice* 28: 221-241.

Watson, Gary. 2004 [1975]. "Free Agency" In G. Watson, *Agency and Answerability: Selected Essays*. New York: Oxford University Press, pp. 13-32.

Watson, Gary. 2004 [1987]. "Responsibility and the Limits of Evil" In G. Watson, *Agency and Answerability: Selected Essays*. New York: Oxford University Press, pp. 219-259.

Watson, Gary. 2004 [1996]. "Two Faces of Responsibility." In G. Watson, *Agency and Answerability: Selected Essays*. New York: Oxford University Press, pp. 260-288.

Watson, Gary. 2011. "The Trouble with Psychopaths," In R. J. Wallace, R. Kumar, S. Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. New York: Oxford University Press, pp. 307-331.

Wolf, Susan. 1987. "Sanity and the Metaphysics of Responsibility." In F. Schoeman (ed.), *Responsibility, Character, and the Emotions: New Essays on Moral Psychology*. Cambridge: Cambridge University Press.

Wolf, Susan. 1990. *Freedom Within Reason*. New York: Oxford University Press.

Wolf, Susan. 2011. "Blame, Italian Style" In R. J. Wallace, R. Kumar, S. Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. New York: Oxford University Press, pp. 332-347.