

## Character and Blame in Hume and Beyond

Antti Kauppinen

Final draft

February 9, 2015

For *Perspectives on Character*. New York: Oxford University Press. Ed. Iskra Fileva

A few months ago, I was at the famously congested and uninspiring Penn Station in New York with a toddler and a baby in a pram, rushing for a train. We were just about to enter an elevator when someone pushed an elderly woman in a wheelchair past us, leaving no room for our pram. Before the elevator doors closed behind the offenders, I lost my precarious cool and uttered a few choice words, and possibly even made a one-finger gesture that may have misleadingly suggested I was a New Yorker myself. There is little doubt that circumstances notwithstanding, swearing at someone with reduced mobility was bad behavior. But was I to blame for it? Suppose, perhaps contrary to fact, that such behavior was entirely uncharacteristic of me – that it didn't in any way indicate that my character is in some way defective. You might then hesitate to condemn *me*, as opposed to my action. And even if you felt somewhat indignant with me, you might not blame me in the sense of regarding our relationship as impaired or withdrawing your good will from me. If so, then at least some important forms of blame seem to involve some sort of link between a bad action and the agent's character.

David Hume, characteristically, takes this line of thought to an extreme conclusion. He maintains that we are *only* to blame for bad actions insofar as they are indications of our character. It is thus the primary focus of the moral sentiments that constitute blame. Hume's claim appears to be in the first instance empirical: contrary to the prejudices of moralists, this is how our actual practice functions. But as he so often does, he seems to move from this

observation to the normative claim that blame is *justified* only insofar our character is defective. On the face of it, this is an implausible thesis, and I don't believe it can be defended without qualification. But there is after all something to say for it, or so I'll argue in this chapter.

I will begin with an outline of Hume's account of blame and its target. Briefly, Hume believes that blame consists in what he calls the indirect passions of hate, contempt, and withdrawal of good will. A person becomes the object of such passions when she performs an action that causes or is apt to cause someone to suffer, and the action is associated with her as a result of issuing from an enduring quality of hers. The blame-constituting passions motivate action to change the agent's character, such as punishment. In the second section, I discuss three challenges to this account. The first is that responsibility requires voluntary control, which we have over actions but not over character traits. The second is that emphasis on character can't account for the difference in our reactions in the cases of moral outcome luck, and the third that we can be blamed for actions that are out of character, and more generally good people can be blameworthy for bad actions. I argue that Hume has more resources to respond to these challenges than it may seem at first sight. Nevertheless, in virtue of such problems, even sympathetic interpreters such as Paul Russell have argued that the strong link between blame and character that Hume posits "must be judged one of the weaker aspects of his general theory of responsibility" (Russell 1995, 102).

In the final section, I situate Hume's view within the contemporary debate, and show that there may be room for an account that emphasizes the importance of character for responsibility. I argue in an ecumenical spirit that the two leading interpretations of blame that have recently emerged can peacefully co-exist. On the one hand, there is reactive blame, which consists in second-personal reactive attitudes that address a demand to the wrongdoer (Strawson 1962, Wallace 1996, McKenna 2012). On the other, there is relational blame,

which consists in a judgment that the wrongdoer's actions have impaired an ethically significant relationship, and consequent modification of intentions and expectations (Hieronymi 2004, Scanlon 2008, Smith 2013). One important piece of evidence for the existence of these two kinds of blame is that there seem to be two kinds of forgiveness corresponding to them. Perhaps the more familiar kind consists, roughly, in ceasing to hold negative reactive attitudes such as resentment or anger towards the person. But one can do so without yet forgiving the culprit in a deeper way, which consists, roughly, in openness to close personal relationships and extending good will towards the other – that is, giving up relational blame. I believe Hume's view is most charitably interpreted as a form of the latter. If its ambitions are restricted in this way, a character-focused account of blame becomes much more plausible.

### **1. Hume on Blame and Character**

Blame is a distinctive sort of response to wrongdoing. It targets the agent, and not the action itself, and thus requires some link between the action and the agent that explains how we can extend our condemnation of the former to the latter (Sher 2006). It has a specific kind of force, which distinguishes it from merely registering that there is something bad about the agent or action. It is linked to or perhaps constitutive of holding the agent responsible or accountable for the action. Consequently, it is not appropriate towards some agents in some circumstances. It may also be inappropriate for some subjects to blame even a responsible agent, and how it is appropriate to blame or express blame may vary depending on the subject's relation to the agent.

A philosophical account of blame should make sense of these platitudes. It is common to hold that some proposals don't go very far. Blame is not just a *judgment* that the agent has a bad character, for we seem to be able to make such judgments without blaming

the person. A judgment or evaluation account doesn't capture the distinctive force of blame. Nor can it be identified with overt *sanctioning or punishing behavior*, since we can blame someone privately. It seems to lie somewhere in between the two.

### *1.1 The Character Thesis and the Nature of Blame*

How does Hume address these desiderata? As Paul Russell (1995) observes, Hume's method is not traditional conceptual analysis. Instead, rather like Peter Strawson (1962), Hume begins with our actual practices of holding each other accountable. Above all, he is interested in the point and nature of holding people responsible and the kind of excuses that we accept, and claims that we don't blame each other for bad behavior (or praise for good behavior) unless we take the action to have issued from the agent's character. The core thesis he defends is thus something like the following:

#### The Character Thesis (CT)

Blame targets a person's character, as manifested by bad thoughts, words, and actions.

Hume frequently contrasts the evaluation of actions with blame towards the agent. For example, he says that while actions are good, perhaps best evidence of character, "'tis only so far as they are such indications, that they are attended with love or hatred, praise or blame" (T 575). He takes very seriously the challenge of connecting the action to the agent.

As he puts it,

Actions are by their very nature temporary and perishing; and where they proceed not from some cause in the character and disposition of the person, who perform'd them, they infix not themselves upon him, and can neither redound to his honour, if good, nor infamy, if evil. The action itself may be blameable. ... But the person is not

responsible for it; and as it proceeded from nothing in him, that is durable or constant, and leaves nothing of that nature behind it, 'tis impossible he can, upon its account, become the object of punishment or vengeance. (T 411)

The next question is, then, why the object of blame and praise should be something “durable and constant”. The answer lies in Hume’s account of blame and praise as indirect passions, and his view of their forward-looking point. In general terms, Hume thinks of blame as a form of moral disapprobation. This has led many interpreters to believe he endorses a purely *evaluative* conception of blame. This would be a problem, since it seems we can judge someone to be bad or flawed without thereby blaming them.

However, on closer inspection, it is clear that by “approbation” Hume doesn’t mean merely evaluation. It is true that in some places, Hume seems to identify approbation with a special kind of pleasure. For example, he says that “the distinguishing impressions, by which moral good or evil is known, are nothing but *particular* pains or pleasures” (T 471). This may seem like a purely evaluative feeling. But elsewhere, Hume clearly takes approbation to consist in what he calls an *indirect passion* that results, in part, from such evaluative feelings. An indirect passion is a passion that arises from an association with a prior impression of pleasure or pain and the idea of its cause. For an example of this “double relation of ideas and impressions”, consider pride: “Any thing, that gives a pleasant sensation, and is related to self, excites the passion of pride, which is also agreeable, and has self for its object” (T 288). So, the idea of a comfortable house gives an independent feeling of pleasure, which is associated (by phenomenal resemblance) with the pleasant feeling of pride, when the idea of the comfortable house is associated somehow with my idea of myself (maybe I built it). These associations between the two impressions and the two ideas together give rise to pride.

For Hume, the second- or third-personal counterpart of pride is love or esteem, which arises from pleasure caused by an enduring quality of another. Similarly, the qualities of others that give rise to pain of the right kind arouse hate or contempt towards the person who is the source of the pain. Love and hate, in turn, give rise to further passions, such as desire to cause happiness or misery to their object, and thus motivate action. While love and hate are not as such moral sentiments (at least not in the relatively narrow contemporary sense), Hume does maintain that approbation and blame are specific forms of love and hate:

The pain or pleasure, which arises from the general survey or view of any action or quality of the mind ... gives rise to our approbation or blame, which is nothing but a fainter and more imperceptible love or hatred. (T 614)

It would be uncharitable to interpret Hume as claiming that to blame someone is to hate them in the ordinary sense, or that we love in the narrow sense everyone we approve. As Páll Árdal says of the positive case, “‘Love’ comes close to being treated as the name for a general favorable attitude to people other than oneself” (Árdal 1977, 414). Indeed, Hume says that in order to keep things simple and avoid “superfluous speculations”, he is going to refrain from “entering into all the minute differences of sentiment” that are excited by qualities we approve or disapprove, and stick to broad categories (EPM 291n2). Nevertheless, he maintains that love in his broad sense manifests itself in “tenderness, friendship, intimacy, esteem, good-will”, among other things. He is not as explicit about the relevant forms of hate, but often pairs hate with contempt and occasionally with disgust. Presumably blame also involves withdrawal of tenderness, friendship, intimacy, and good-will, at least to a degree. These indirect passions of approbation and disapprobation are often what Hume calls “calm passions”, which are “desires and tendencies, which, tho’ they be real passions, produce little emotion in the mind, and are more known by their effects than

by the immediate feeling or sensation” (T 417). A Humean passion thus need not continuously manifest itself in our consciousness, and blame can take the form of a seemingly non-emotional tendency to act in certain ways towards the agent, such as avoiding their company.

So, on this reading, blame is an indirect passion of hate or contempt directed towards a person as a result of associating her with an action that is painful. In order for us to associate an action with a person, we must take it to have issued from something in her, and not just linked to her by chance. This is the first reason why only actions that issue from someone’s character give rise to blame. Otherwise they would not be associated with the person herself, and wouldn’t give rise to the indirect passion that constitutes blame.

The second reason, as I see it, is linked with blame’s forward-looking function. When we express blame by publicly denouncing or sanctioning someone, we may dissuade them from performing bad actions in the future, and perhaps even change their dispositions so that they’re less likely to do so in the absence of external control. This evidently makes little sense in the case of someone whose bad action didn’t reflect an enduring trait in the first place, and thus wasn’t going to repeat the behavior in any case. Hume seems to be gesturing towards this when he says, in the passage already quoted, that unless an action proceeds from a “durable and constant” trait of an agent, “tis impossible he can, upon its account, become the object of punishment or vengeance.” (T 411) Some readers of Hume, such as Philippa Foot (1966), have overemphasized this aspect of his view, and taken him to hold a purely utilitarian view of blame. As we’ve already seen, that’s not the case. It is more plausible to take him to have held that forward-looking considerations provide additional justification for the practice of blaming people on the basis of character traits manifest in action (Russell 1995). The insufficiency of the utilitarian interpretation can also be observed

in Hume's belief that it can be appropriate to blame people even when doing so will effect no change in their character, for example when they're already dead (T 584).

The final kind of evidence that Hume offers for CT is found in his discussion of *excuses*. His claim is that what unifies commonly accepted excuses is that they show the bad action does not issue from an enduring trait of the agent. One common excuse is that the harm caused by an action was not intended by the agent, but was only accidental. Here Hume observes that "as the character of a person is no wise interested in such injuries as are casual [i.e. accidental] and involuntary, it seldom happens that on their account, we entertain a lasting enmity" (T 350). In other words, lack of intention works as an excuse (to the extent it does), because it indicates that the particular action doesn't reflect who the agent is. Similarly, what explains why ignorance is an excusing condition (when it is such) is that if I perform an action due to a mistaken belief about the nature of its aim or means, "[n]o one can ever regard such errors as a defect in my moral character." (T 460)

Conversely, when one does intend to do harm, this reveals something about their character: "An intention shews certain qualities, which remaining after the action is perform'd, connect it with the person, and facilitate the transition of ideas from one to the other" (T 349). However, since it is ultimately the link to character that matters, the presence of intention is not in all cases necessary for blameworthiness: "If that quality in another, which pleases or displeases, be constant and inherent in his person and character, it will cause love or hatred independent of the intention" (T 348). In Hume's example, a person's folly will give rise to aversion, even though they don't mean to do foolish things. So the absence of intention is not a foolproof excuse, since intention is not always necessary to connect an action with the agent.

## 1.2 Hume on Character



So far, I've been discussing why Humean blame attaches only to actions that issue from someone's character. But what precisely is character, for him? Clearly, in order to fit the role just sketched, it must be an *enduring* feature of a person that plays a significant role in causing actions. But which enduring features of a person form their character? In the background of Hume's idiosyncratic view lies his skepticism concerning the self and causal powers, which sit uneasily with the commonsense conception of character traits as dispositions constitutive of who we are. As a dedicated empiricist, Hume sets a high bar for ontological commitment. Regarding the self, all that experience warrants is postulating a "succession of related ideas and impressions, of which we have an intimate memory and consciousness" (T 277). When it comes to powers or dispositions, he maintains that "the distinction, which we often make betwixt power and the exercise of it is ... without foundation" (T 171).

How can these views be reconciled with the rather commonsensical use that Hume makes of the notion of character in his moral philosophy? If character traits are elements of the self, they must consist in ideas or impressions. Many interpreters of Hume have indeed taken him to regard character traits as *passions*, which are a kind of impression. But how could a passion – rather than a disposition to feel in certain ways, say – be a character trait? Paul Russell (1995) draws attention to Hume saying that a passion can become "a settled principle of action" or even "the predominant inclination of the soul" (T 419). On his reading, character traits are thus enduring passions, presumably calm ones, which result in dispositions to act in certain ways, but also manifest themselves in other ways, such as generating other passions. Hume does sometimes refer to ambition, avarice, and vanity, for example, as passions (EPM 259). On the basis of this sort of reasoning, Jane McIntyre concludes that "on a Humean view, character is the structured set of relatively stable passions that give rise to a person's actions" (McIntyre 1990, 201).

However, any reading that identifies character traits as a kind of passion has a forced quality, and while it is certainly motivated by a broader view of Hume’s writings, it does not fit with other things he says about character. For Hume also considers other mental qualities, such as good sense, judgment, and wit, as virtues or constituents of personal merit, and *a fortiori* as character traits in the relevant sense. In the *Treatise*, he call such traits “natural abilities”, and in the second *Enquiry* “talents”. They are mental qualities just as social virtues are, and “equally produce pleasure; and of course have an equal tendency to procure the love and esteem of mankind” (T 606–607). Qualities or habits such as “indolence, negligence, want of order and method, obstinacy, fickleness, rashness, credulity” that reduce someone’s capacity for “business and action” are “instantly blamed, and ranked among his faults and imperfections” (EPM 232). Like social virtues, intellectual virtues such as prudence and discernment influence our conduct (EPM 289). Yet these traits, along with industry, frugality, and perseverance, don’t involve any “immediate sentiment”, and are only known by their effects, as any other dispositions are. So it seems clear that Hume takes character to comprise traits that cannot plausibly be interpreted to be passions (or any other kind of impressions or ideas).<sup>1</sup>

If character traits are not passions or perceptions in the Humean sense, they cannot be parts of the self in the Humean sense. Yet they do seem to define who we are, and are the objects of pride or blame on Hume’s own account. We might have to acknowledge that Hume’s sensible observations when he’s doing moral philosophy can’t easily be reconciled with his empiricist epistemology and metaphysics. Alternatively, a Humean might distinguish between a *metaphysical* notion of the self as the succession of perceptions that underlies our identity through time, and a *psychological* notion of the self as a set of traits

---

<sup>1</sup> Mahoney (2009) offers evidence that habits and consciously adopted general rules also form a part of character for Hume.

<sup>2</sup> This was suggested to me by Iskra Fileva in correspondence.

<sup>3</sup> Since Hume appeals to our actual practice, it is worth pointing out that recent studies suggest that ordinary

and states that defines who we are at the moment, relegating character to the latter category.<sup>2</sup> That would mean that our character would only be contingently associated with who we are, metaphysically speaking, which would not be an entirely happy outcome either.

## 2. Challenges to Hume's Account

There is an obvious challenge to any view for which the proper object of blame is character, especially in the inclusive sense that Hume uses: we don't choose our character traits, and choice seems to be linked with responsibility. More precisely, the argument is the following:

1. What we are to blame for must be something that is under our voluntary control.
2. Our character traits are not under our voluntary control.
3. Therefore, we are not to blame for our character traits.

While many others try to finesse the second premise, Hume's main response to this argument is to reject the first. He argues that it rests on a mistaken conception of the significance of voluntariness, and conflicts with our actual practice. He also offers an error theory to account for why people endorse it.

First, on Hume's account of moral responsibility, what fundamentally matters not whether our actions are voluntary but whether they reflect who we are. On his view, our actions are not in any sense undetermined but instead necessitated (in his constant and invariant conjunction sense of necessity) by our "motives, tempers, and circumstances" (T 401). This is something we in fact attribute to others, and rely on in everyday practice, but find it hard to acknowledge in our own case, since the idea of necessity seems to "imply something of force, and violence, and constraint" (T 407). But that is a mistake, since causation is not compulsion. It does not follow from such necessitation that we are not free

---

<sup>2</sup> This was suggested to me by Iskra Fileva in correspondence.

or responsible. Indeed, if we had the ‘liberty of indifference’ – if our will or actions were uncaused and undetermined by our character – we could not be held responsible, since there would be nothing to connect *us* with the action resulting from some chance event (see also Frankfurt 1988). What free will and responsibility require is simply ‘liberty of spontaneity’: our actions are caused by our will, rather than some external factor such as force or violence.

However, the significance of being caused our will turns out to be derivative. When our will doesn’t reflect our character, such as when it results from invincible ignorance of fact, we are not responsible for the action. And as Hume’s discussion of excuses shows, he maintains that in actual practice, we blame people for unchosen traits. He says that “[I]f we can find any quality in a person, which renders him incommodious to those, who live and converse with him, we always allow it to be a fault or blemish, without any farther examination.” (T 606) (Whether finding something to be a ‘fault’ or ‘blemish’ suffices for blame is an issue we need to return to.) Fundamentally, for a Humean compatibilist, our actions are free when they issue from ourselves – when they are necessitated by certain of our own mental qualities.<sup>3</sup> Somewhat ironically, Hume is thus a proponent of what is often called a ‘deep self’ view of responsibility (Wolf 1987, Sripada MS). Voluntary control matters derivatively, because it *typically* indicates that the action manifests our character, or who we are. And if voluntariness is of strictly derivative relevance and what is fundamental is that an action reflects our character traits, it is no objection to blaming someone for a character trait that it is not under voluntary control.

Finally, Hume offers an error theory of why people have thought that voluntariness is significant for responsibility. The reason is that our will, and character traits we can shape through voluntary practice, are susceptible “to be chang’d by the motives of reward and

---

<sup>3</sup> Since Hume appeals to our actual practice, it is worth pointing out that recent studies suggest that ordinary people seem to implicitly endorse some form of compatibilism (Murray and Nahmias 2014). That is, they are willing to blame people for bad actions, as long as the actions issue from people’s own choices and deliberation, even if determinism is true.

punishment, praise and blame” (T 609), while dispositions like prudence are not. Moralists and legislators with forward-looking concerns have thus focused on praising and blaming only what is under voluntary control. But ordinary people who lack such aims of improving society “naturally praise or blame whatever pleases or displeases them”, regardless of whether the quality is susceptible to change. Since the warrant for blame doesn’t hang on its good consequences, it is the moralists focusing on voluntariness who are in error.

Other challenges to Hume may be more damaging. One issue concerns *moral luck*. Hume clearly believes that constitutive moral luck, turning out to be a certain kind of person due to factors beyond one’s control, is compatible with being blameworthy, as we have just seen. The challenge is that he also holds that *consequential* moral luck, what our actions turn out to be due to uncontrollable factors, is morally irrelevant. To take Adam Smith’s example, two people may both recklessly throw a stone over a wall; one lands harmlessly, while the other kills a man (TMS 120–121). It is a fact about our actual practice that at least when it comes to reckless endangerment, we blame (and punish) people more depending on the actual consequences of their actions, even if they are due to sheer luck. The man whose stone lands harmlessly isn’t brought to the scaffold. Hume refuses to acknowledge this. He accepts that when “good disposition is attended with good fortune”, it gives us stronger pleasure to observe it, but nevertheless, “we do not say it is more virtuous, or that we esteem it more” (T 585). Consistency requires him to say the same in the case of blame, thus preserving CT. But as an empirical claim about our practice, it is false.

The argument can be put in the following terms:

1. Two actions that issue from identical character traits may have consequences that differ greatly in value due to a consequential moral luck.
2. We blame people who have bad consequential moral luck, while we don’t blame people who have good consequential moral luck (or don’t blame them as much).

3. Hence, our actual blaming practices do not just depend on character traits manifest by actions, but also consequential moral luck.

It is in principle open for Hume to say that our practices are mistaken: insofar as two people's actions manifest the same character traits, they're equally blameworthy, even if only one of them happens to, say, cause the death of a child. But this is hard to reconcile with the spirit and letter of Hume's practice-based approach, given that we do actually seem to praise and blame people more on the basis of consequential moral luck. So it seems something has to give. The only consolation for CT is that moral outcome luck is puzzling on most other accounts of blameworthiness as well. I will set this issue aside in what follows.

The third critical argument is straightforward. We sometimes blame people for actions that do not manifest vice or are otherwise "out of character" for them. For example, it is surely possible for a kind person to say a mean thing when stuck at Penn Station at rush hour. After all, it doesn't suffice for one to have a vice or a disposition to do some bad thing X if one does X once (or a few times), since character traits are or involve dispositions to act, feel, and desire in certain ways across a variety of situations. Nevertheless, we may blame the person who says something mean under pressure. So CT seems to be false.

There are two ways a Humean might respond to this argument. The concessive response would be to say that while we don't, and shouldn't, blame people for out of character actions *in the Humean way*, there are other kinds of blaming response that are appropriate in such cases. I will discuss this possibility in the next section. The other response is to maintain that on a more fine-grained conception of character, seemingly out of character actions are not such after all. The fact that someone says a mean thing when they're having a bad day shows that she "always had it in her" to act that way, while a kinder person wouldn't have been mean even when having a bad day. It's only actions that result

from excusing conditions like force or ignorance that are genuinely out of character, because they don't reflect durable principles of the mind that are features of the individual. Relatedly, in her contribution to this volume, Nomy Arpaly points out that we sometimes regard drunken actions, such as a married person's kissing of a co-worker at a party, as revealing the person's true self. On the Humean view, this is the case when the loss of self-control resulting from inebriation causes an otherwise suppressed promiscuous disposition to manifest itself. If, in contrast, the agent had no underlying inclination to kiss her co-worker that she was holding back in her sober state, the action may be genuinely out of character, and the Humean predicts we attribute it to the drink and not the person, at least mitigating the blame.

George Sher (2006) considers the second type of response, but argues that it overgeneralizes. Pretty much anyone has dispositions such that they would in some circumstances say something mean or keep something that is not theirs. It can't be relevant to blameworthiness if you have a complex set of dispositions that results in your being prone to meanness at Penn Station, while someone else has a tendency to do the same sort of things at natural parks. As Sher puts it, "I suspect that for (just about) every person and (just about) every type of bad act, there is some conceivable set of conditions under which that person would perform an act of that type." (2006, 26) So his worry is that we all turn out to be (perhaps equally) blameworthy, even if we have had the circumstantial moral luck not to end up in a situation in which we would have acted badly. But this is absurd. Good people can do bad things and be blameworthy for it, but "we cannot appropriately blame a good (though imperfect) person if he does not act badly" (Sher 2006, 28). Even *bad* people are to blame for bad acts, not for their bad character (at least not exclusively).

Sher is no doubt right that we can't be blamed for what we might have done (though what we would have done in someone else's shoes may be relevant to our *standing* to blame).

But as I've formulated CT, it doesn't say we're to blame for bad character as such, but only bad character that is manifested in actions that issue from our character (whether or not the character is bad). It is a matter of circumstantial moral luck, to be sure, whether we reveal our character in action. If Penn Station is demolished, some people may never say mean things, because their disposition to do so is never triggered. But it is a consistent Humean view that such luck doesn't undermine the justification for blame, when we *do* end up doing bad things. If constitutive bad luck doesn't undermine blameworthiness, why should circumstantial bad luck do so, as long as our actions still reflect who we are?

### **3. Two Kinds of Blame?**

In the last section, I outlined three major challenges for Hume's Character Thesis: sometimes he speaks as if blaming was just negatively evaluating someone, he doesn't acknowledge the role of consequential moral luck, and struggles to account for blame for uncharacteristic actions. These issues are not unrelated. If blaming someone was just a matter of thinking that they are morally defective, it would make a good deal of sense that we wouldn't blame people for bad consequences beyond their control or uncharacteristic actions, since they don't necessarily indicate that the person is morally defective. Hume certainly flirts with such a purely evaluative view. For example, he says that "A blemish, a fault, a vice, a crime; these expressions seem to denote different degrees of censure and disapprobation; which are, however, all of them, at the bottom, pretty nearly of the same kind or species." (EPM 295)

Alas, there does seem to be more to blame than evaluation – it has to have a "quality of opprobrium" lacking in our responses to, say, defects of physical appearance, as Jay Wallace emphasizes. And as I've argued, Hume does seem to recognize this in discussing blame as an indirect passion. Still, the sentiments he talks about, such as hate and disgust, do



not play a major role in contemporary sentimental accounts of blame. So should we relegate the Humean view to being a historical curiosity? In this final section, I argue that if we adopt a broader conception of blame, there is after all something to be said for Hume's account, and the way it links blame to character. More specifically, I'll sketch two conceptions of blame that have emerged in contemporary debates, and argue that they can co-exist. Interestingly enough, Hume's and Adam Smith's views can be seen as complementary sentimental accounts of different varieties of blame.

### *3.1 Reactive and Relational Blame*

The first variety of blame I will call *reactive blame*. As the name indicates, reactive blame consists of reactive attitudes such as resentment, indignation, and guilt. Their target, in Strawsonian terms, is an *undesirable quality of the will* manifest by an agent's action. As Adam Smith puts it, "To the intention or affection of the heart ... all praise or blame, all approbation or disapprobation, of any kind, which can justly be bestowed upon any action, must ultimately belong." (TMS 109) (Note that Smith thus rejects CT.) When someone acts in a way that shows disrespect, neglect, or malice towards us, we respond with a negative attitude of our own. As Stephen Darwall (2006) nicely observes, reactive attitudes do not just motivate us to sanction or punish the agent, but *address* the agent – they come with an "RSVP", as he puts it. We expect the agent to acknowledge the claim that we, or someone else, had on their behavior. This is one sense in which reactive attitudes are second-personal rather than third-personal or objective: they call on the agent to respond, and thus assume them to be capable of doing so. In Smith's words,

- . The object ... which resentment is chiefly intent upon, is not so much to make our enemy feel pain in his turn, as to make him conscious that he feels it upon account of

his past conduct, to make him repent of that conduct, and to make him sensible, that the person whom he injured did not deserve to be treated in that manner. (TMS 112)

The RSVP for resentment, thus, is guilt and repentance, followed by actions that distinctively manifest them, such as repentance, apology, and making amends. Other reactive attitudes, such as contempt, call for a different response, such as shame on the part of the agent. When the agent does respond in such ways, it may be appropriate for the victim to *forgive* her – to moderate or give up negative reactive attitudes.

Since reactive attitudes target the quality of the will with which the agent acted and not an enduring disposition or trait of the agent, a proponent of a reactive blame account owes us an explanation of why it is proper to target the agent for a mental state they quite possibly no longer have. One promising answer is offered by Pamela Hieronymi, who observes that past wrongs which haven't been apologized or atoned for can be seen as making the claim that such treatment of the victim is acceptable (2001, 546). Unless the agent, and often others, respond to the address, the implicit false claim regarding the victim's status is left to stand.

Assuming a response along these lines works, the main outstanding critique of reactive attitude accounts of blame is that reactive attitudes are not *necessary* for blame. For example, many argue that we can blame distant or historical actors without feeling resentment or indignation towards them, or addressing them in any way (e.g. Scanlon 2013). I cannot here explore possible responses to this point. In any case, for my pluralist proposal, it is not a problem, as long as the non-reactive cases of blame fall under the second category of blame that I allow.

The second kind of blame I will call *relational blame*. Thomas Scanlon, above all, as recently defended the idea of seeing blame in terms of the relationships we have to one

another. Roughly, he believes that to regard someone as blameworthy is to take them to have attitudes that impair the relations we can have with them, and to blame them is to regard them as blameworthy and modify the relationship accordingly (Scanlon 2008, 128).

Relationships, for Scanlon, are constituted by attitudes (which are in turn largely a matter of what one takes to be a reason for or against something), intentions, and dispositions to behave. On this broad view of relationships, even strangers stand in a moral relationship to each other, though Scanlon's claim is most persuasive in the case of specific relationships (Sher 2013). For example, for us to be friends, we must have a certain good-will towards each other, take pleasure in each other's company, be disposed to spend time together and make choices that benefit the other beyond the call of duty, among other things.

But what if it turns out you make fun of me behind my back? I may believe that given the nature of our relationship and the standard for attitudes that is implicit in it, your willingness to do what you've done gives me reasons to change my attitudes and dispositions (this is the blameworthiness judgment) and as a result withdraw my good-will and lose the desire to spend time with you. I may no longer trust you, or take pleasure in your success. I might also, but need not, resent you. Still, Scanlon maintains, I would be blaming you, were I to revise my attitudes and intentions as described, as a result of believing that you've impaired our relationship.

These attitudes do not seem to be second-personal in the sense that reactive attitudes are. When I cease to take pleasure in your success, given your mean treatment of someone else, I don't in any sense address you. My attitude change doesn't come with an RSVP, in Darwall's terms. Indeed, Scanlon explicitly says that "blame itself – the revision of one's attitudes toward a person in response to attitudes expressed in his behavior – is not, even incipiently, a form of communication" (2008, n54). But if there's no moral address, can our response still amount to the sort of condemnation that is involved in blaming someone?

Scanlon maintains “If I “write someone off ” as a person I am going to have nothing to do with, then I am blaming him, even if this is accompanied by no hostile feelings, perhaps because I regard him as not worth being angry at.” (2008, 160) After all, being written off by someone in virtue of what we’ve done is not pleasant or desirable for most of us. So there is certainly an aspect of condemnation in relational blame.

Reactive and relational accounts of blame are standardly presented as competing views of the same phenomenon. They are clearly distinct along the dimensions of opprobrium and second-personality. But must we see them as competitors, or do we have sufficient reason to conclude that there are two kinds of genuine blame? As I’ve suggested, I think that we should take the ecumenical option. One good reason to do so is that there seem to be two kinds of *forgiveness*, that is, two ways of ceasing to blame someone. Suppose that a husband has a brief affair with an acquaintance, and his wife finds out about it. At first, let us stipulate, she is angry, even indignant with him, blaming him for wronging her by violating her trust. Marriages being complex, they nevertheless stay together, and he shows himself to be deeply sorry for what he has done. Gradually, she ceases to resent him. In one sense, she has forgiven him. Yet the absence of negative reactive attitudes doesn’t mean that their relationship has returned to what it was at its best. There may still be a lack of trust and joy, and the plans they make together may center around the children rather than romantic adventures for two. Perhaps, however, after there’s a sufficient amount of water under the bridge, this, too, will change. The wariness is gone, and they can laugh together at what has happened. The relationship is no longer marred by the husband’s past infidelity. At this point, it makes sense to say that the wife has forgiven the infraction in a deeper way. And if the point never comes, it may be that after years of faithfulness and co-operation, the wife’s sister may legitimately ask why she still blames him.

If there are these two kinds of forgiveness – call them *attitudinal* and *relational* forgiveness – and forgiveness is a matter of ceasing to blame, we have good reason to think there are two kinds of blame, too. These are obviously the reactive and relational kinds. The ecumenical view allows us to say that someone can relationally blame someone without reactively blaming them – having seen you being rude to someone, I may not resent you, but I will not invite you to my party. Someone may also reactively blame someone without relationally blaming them – I can resent you for what you just said, but I won't let it impair our relationship, since it's so important to me. Being able to capture the nuances and distinctions involved in blame and forgiveness strongly favors a hybrid account.

### *3.2 Character and Relational Blame*

What does the distinction between reactive and relational blame mean for CT and Hume's account in general? As I already observed when discussing the difference between Hume and Smith, Humean blame does not consist in reactive attitudes. So it may be best seen as an account of relational blame. And in fact, what he says about the indirect passions of hate and contempt fits fairly well in the picture that Scanlon sketches. After all, Humean blame is a calm passion that is manifest in tendency to avoid doing certain things with the object of blame, and rules out tenderness, intimacy, and good will. Insofar as relationships supervene on attitudes and intentions, as Scanlon says, Humean blame does amount to modification of a relationship.

Of course, for Scanlon such modification of attitudes and intentions only amounts to blame if it results from a judgment that the target's own attitudes have already impaired the relationship. For Pamela Hieronymi (2004), a judgment to the effect that the agent has shown disregard or disrespect already amounts to blame. In Hume's case, it is controversial whether any moral judgment is distinct from the attitudinal response. But it does seem to be

his view that thinking that someone is morally blameworthy is a belief about the tendency of the character traits manifest in the agent's action to arouse blame in anyone who is acquainted with her and regards her from the "common point of view", a perspective people with normal emotional dispositions can share (see Cohon 2008). Sentiments arising from such calm, impartial consideration are what we implicitly appeal to when we use of moral language, and "can alone be the foundation of morals, or of any general system of blame or praise" (EPM 260). So Hume need not claim that just any loving or hating response to someone amounts to approbation or blame. At least, he maintains that our initial responses are to some degree amenable to correction, and it is open to him to say that a sentiment that doesn't receive support from a belief about responses from a common point of view doesn't amount to full-fledged blame.

What happens to the Character Thesis, if we adopt this conception of Humean blame? We can first reformulate it in more modest terms:

*The Relational Character Thesis (RCT)*

Relational blame targets a person's character, as manifested by bad thoughts, words, and actions.

How does RCT fare with the challenges to CT, leaving moral outcome luck aside? First, it certainly alleviates concerns with responsibility. When we find out that a difficult person or even a criminal is the product of a childhood marred by alcohol and violence, we may continue to regard their character as bad, but fairness may require reducing our resentment. But even if reactive blame is inappropriate towards someone who has grown up twisted, relational blame may not be. Withdrawing our affection and intimacy, perhaps even a kind of contempt may be appropriate towards someone who habitually shows disregard for the claims of others.

Second, what about actions that are out of character in the everyday, coarse-grained sense? Take the generally kind person who makes a cutting remark in particularly stressful circumstances, thereby hurting someone. In the circumstances, they took there to be sufficient reason to make a comment about someone recent change of appearance, say, although in a vast majority of similar situations such a thing would never occur to them. Still, given the quality of the will, or the meaning of the action, to use Scanlon's terms, reactive blame may be appropriate towards the person. But what about relational blame? For Scanlon, it seems, it should be appropriate, since he takes it to be based on the attitudes displayed by the agent, regardless of whether they are rooted in her character. RCT, however, says that relational blame isn't appropriate. I believe the Humean view offers a better interpretation of such scenarios. When we regard a bad action as an isolated incident that doesn't reflect who someone is, we do not (and should not) modify our relationship with them. I'll continue to trust and wish the best for a friend who has a laugh at my expense in a specific situation, unless I take it to reveal something hitherto unsuspected about their character and dispositions towards me. In brief, it seems our relationships are not impaired by bad actions that are regarded as being out of character, as RCT predicts.

### *3.3 Conclusion*

In this final section, I've argued that blame comes in both reactive and relational varieties, and that Hume's view is most charitably interpreted as being a form the latter. If this is right, it is much more plausible that a kind of blame is warranted only for actions that reveal our underlying character. After all, when we genuinely think that an action fails to express who the agent really is, it is much less likely to affect how we relate to her. On this interpretation of blame, the crucial question for a Humean account turns out to be whether character is

central to the relationships we have with each other. To the extent that it is, the Relational Character Thesis has more going for it than first meets the eye.<sup>4</sup>

## References

- Árdal, Páll. Another Look at Hume's Account of Moral Evaluation. *Journal of the History of Philosophy* 15 (4), 405–421.
- Coates, Justin and Tognazzini, Neal 2013. *Blame: Its Nature and Norms*. New York: Oxford University Press.
- Cohon, Rachel 2008. *Hume's Morality: Feeling and Fabrication*. Oxford: Oxford University Press.
- Darwall, Stephen 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- Foot, Philippa 1966. Free Will as Involving Determinism. Reprinted in *Virtues and Vices*. Oxford: Oxford University Press 2002, 62–73.
- Frankfurt, Harry 1988. Rationality and the Unthinkable. In *The Importance of What We Care About*. Cambridge: Cambridge University Press.
- Hieronymi, Pamela 2001. Articulating an Uncompromising Forgiveness. *Philosophy and Phenomenological Research* 62 (3), 529–555.
- Hieronymi, Pamela 2004. The Force and Fairness of Blame. *Philosophical Perspectives* 18, 115–148.
- Hume, David. 1739–40/1978. *A Treatise of Human Nature* (T). Ed. L.A. Selby-Bigge and P. H. Nidditch. Oxford: Oxford University Press.

---

<sup>4</sup> I want to thank Iskra Fileva, Lilian O'Brien, and participants in my moral sentimentalism seminar at the University of Jyväskylä in September 2014, especially Aino Lahdenranta and Ninni Suni, for feedback and ideas.



- Hume, David. *An Enquiry Concerning the Principles of Morals* (EPM). In Geoffrey Sayre-McCord (ed.) *Moral Philosophy*. Indianapolis: Hackett.
- McIntyre, Jane 1990. Character: A Humean Account. *History of Philosophy Quarterly* 7 (2), 193–206.
- McKenna, Michael 2012. *Conversation and Responsibility*. New York: Oxford University Press.
- Mahoney, Robert Heath 2009. *Hume's Conception of Character*. PhD Dissertation, University of Southampton.
- Murray, Dylan and Nahmias, Eddy 2014. Explaining Away Incompatibilist Intuitions. *Philosophy and Phenomenological Research* 88 (2), 434–467.
- Russell, Paul 1995. *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility*. Oxford: Oxford University Press.
- Scanlon, Thomas 2008. *Moral Dimensions: Permissibility, Meaning, and Blame*. Cambridge, MA: Harvard University Press.
- Scanlon, Thomas 2013. Interpreting Blame. In Coates and Tognazzini (eds.) 2013, 84–99.
- Sher, George 2006. *In Praise of Blame*. New York: Oxford University Press.
- Sher, George 2013. Wrongdoing and Relationships: The Problem of the Stranger. In Coates and Tognazzini (eds.) 2013, 49–65.
- Smith, Adam 1759–90/2002. *The Theory of Moral Sentiments* (TMS). Ed. Knut Haakonssen. Cambridge: Cambridge University Press.
- Smith, Angela M. 2013. Moral Blame and Moral Protest. In Coates and Tognazzini (eds.) 2013, 27–48.
- Sripada, Chandra MS. Self-Expression: A Deep Self Theory of Moral Responsibility. Online at <https://umich.app.box.com/s/qdm2tcwekrjtz6co4ob1>. Accessed January 20, 2015.

Strawson, Peter. 1962. Freedom and Resentment. *Proceedings of the British Academy* 48, 1–25.

Wallace, R. Jay 1996. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

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

Wolf, Susan 1987. Sanity and the Metaphysics of Responsibility. In Ferdinand Schoeman (ed.), *Responsibility, Character, and the Emotions*. Cambridge: Cambridge University Press, 46–62.