



THE GOOD, THE BAD AND THE BLAMEWORTHY

BY NEIL LEVY

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

VOL. 1, No. 2 | JUNE 2005

URL: WWW.JESP.ORG

COPYRIGHT © NEIL LEVY2005

The Good, the Bad and the Blameworthy

Neil Levy

ACCOUNTS OF MORAL responsibility come in two main flavors. There are accounts that hold that an agent is responsible for something (an act, omission, attitude, and so on) just in case that agent has – directly or indirectly – *chosen* that thing, and there are accounts that hold that an agent is responsible for something just in case that thing is appropriately *attributable* to her. Each kind of account explains many cases well, and each captures a great many of our pretheoretical intuitions about responsibility; each also yields, occasionally, somewhat counterintuitive results. Call these accounts *volitionist* and *attributionist* accounts of moral responsibility. Attributionism has a number of distinguished and able defenders. However, I shall argue, it is wrong: Volitionism is superior, because it alone can accommodate the relatively stringent epistemic conditions that any adequate theory of moral responsibility must recognize, and it alone can accommodate the intuitively powerful distinction between *bad* agents and *blameworthy* agents.

Before outlining the contending accounts of moral responsibility, let me say a few words about what they are accounts *of*. In order to avoid the twin risks of begging the question against either account or simply talking past one another, we need a shared notion of moral responsibility. A fully adequate definition must await the development of a complete theory of moral responsibility, but the following condition upon such a theory will serve to guide our quest: To say that an agent is morally responsible (for an act, omission or attitude) is to say that the Strawsonian reactive attitudes are justified in relation to her with regard to that act, omission or attitude (Strawson 1962). That is, it is appropriate for observers to have certain attitudes in relation to her and her act, especially the attitudes, partly cognitive and partly constituted by emotion, of praise and blame.¹ It is a *further* question whether it would be appropriate to punish or reward the agent for her act, or even whether it would be appropriate to *express* the judgment. It may be that the expression of the reactive attitudes is justified under stronger, or merely different, conditions than those under which it is appropriate merely to have them, and it is with the latter that we are here exclusively concerned.²

¹ If it is appropriate for third persons to have the reactive attitudes with regard to a person, is it also appropriate for the person herself to have the first-person analogs of these attitudes (guilt, pride, shame, and so on)? Perhaps, but I prefer to set this issue aside. Bernard Williams (1981) has famously argued for the rationality of certain first-person reactive attitudes in the absence of responsibility; if this is right, then the first-person attitudes might have quite different conditions to the third-person.

² For instance, the (quasi-) expression of attitudes the *having* of which is not justified might be justified on consequentialist grounds, as indeed might be the imposition of sanctions. But

Attributionist and Volitionist Accounts of Moral Responsibility

In a well-known paper, Gary Watson (2004b) distinguishes what he calls the two faces of responsibility. The first, which he calls the aretaic or attributability aspect of responsibility, is intimately linked to a self-disclosure view of moral responsibility. Someone is responsible, in this sense, if her action is expressive of who she is and where she stands on questions of value. The second face of responsibility Watson calls the accountability aspect. Watson argues that someone is accountable for an action if sanctions (or benefits; from now I shall concentrate on the negative case) are fairly applied to them as a consequence of it. We have defined moral responsibility in such manner as to exclude the question of the appropriateness of sanctions. That issue aside, Watson's distinction seems to map neatly onto the distinction between responsibility as understood by attributionism and responsibility as understood by volitionism.

Whereas an act is attributable to an agent if it is expressive of who she is, agents are accountable for actions only if they had a reasonable opportunity directly or indirectly to avoid infringing the standards for the violation of which they are held responsible (Watson, 2004b: 276). Agents are able directly to avoid such an infringement if they are able to conform their acts to the relevant standard; they are able indirectly to avoid infringement if they were able to avoid being held to that standard at all. Thus, agents can be responsible for their failed actions in the absence of a capacity to conform to the relevant standard just in case they could have avoided the requirement in the first place – for instance, someone who accepts the role of a doctor, knowing what kinds of skills are required to occupy it competently, is not excused responsibility for harming her patients on the grounds that she lacked the skill to do better, so long as she had the opportunity to avoid accepting the role.

Accountability, as Watson understands it, therefore has fairly stringent *control* conditions upon it. But attributability requires relatively little in the way of control. An act or omission can rightfully be attributed to me whether or not I ever exercised control over acquiring the attitude that it expresses. So long as my action is rightly taken to be expressive of my real self – so long, that is, as it is the product, in the right kind of way, of my beliefs and desires, values and commitments, and not of hypnosis, brain manipulation, mental illness or what have you – then it is properly attributable to me.

To be sure, control is not irrelevant on attributionist accounts. Absence of control may block attribution to me of the attitudes apparently expressed by my acts. Suppose that my action was not a product of my beliefs and desires, but instead produced by the intervention of a nefarious neuroscientist, by coercion or by certain kinds of (transient) mental illness. In that case, the

it does not follow, from the putative fact that treating an agent *as if* she were responsible is justified, that the agent *is* responsible.

action does not reflect where I stand on questions of value, and cannot be attributed to me. Absence of control matters only to this extent, on the attributionist account.

For Watson, there are two, mutually irreducible, kinds of responsibility. Volitionists and thoroughgoing attributionists dissent from Watson's claim. Each side believes that there is just one kind of responsibility, though they differ as to which of Watson's notions should be identified with genuine responsibility.

Thoroughgoing, or *strong*, attributionism has been defended by Timothy Scanlon (1998) and Angela Smith (2005), both of whom argue that responsibility-as-attributability is all the responsibility there is.³ To be sure, Scanlon (at least) agrees with Watson that when it comes to imposing sanctions, questions need to be asked about control and avoidability (1998: 286), but these questions do not concern whether a different notion of responsibility applies. Instead, they concern the fairness of imposing a burden upon someone. When we impose such sanctions, we express the only kind of responsibility there is: responsibility-as-attributability.

On the attributionist account, I am responsible for my attitudes, and my acts and omissions insofar as they express my attitudes, in all cases in which my attributes express my identity as a practical agent. Attitudes are thus expressive of who I am if they belong to the class of *judgment-sensitive attitudes*. Judgment-sensitive attitudes are attitudes that, in ideally rational agents, are sensitive to reasons, such that these agents have them when, and only when, they judge there to be sufficient reason for them (Scanlon 1998: 20). Insofar as we are rational agents, we are not simply "stuck" with our judgment-sensitive attitudes. Instead, they are the product and the expression of ourselves as agents. We can therefore appropriately be asked to justify them. It makes no sense to ask me to justify my height, my skin color or my compulsions, simply because none of these aspects of me are sensitive to my judgments. But I can be asked to justify my political views, my fundamental values and my sense of what is important and what trivial. My judgment-sensitive attitudes reveal where I stand on questions of value.

Assessing the Attributionist Claim

Is attributability really (a kind of) responsibility? That is, are the reactive attitudes, especially praise and blame (and perhaps their first-person equivalents) really justified whenever we can rightfully attribute a morally relevant justification-sensitive attitude to an agent? I shall argue that they are not, and that therefore attributability is not moral responsibility. Moral responsibility has much stronger control conditions than attributionism can accommodate, including, importantly, epistemic conditions. No such

³ Strong attributionism has also been defended by Adams (1985), and perhaps Greenspan (2001), though her view retains volitionist elements.

conditions need be satisfied for an attitude to be attributed to an agent (as attributionists recognize), but when they are not satisfied the reactive attitudes are unjustified.

I shall not present much in the way of argument for the claim that moral responsibility has stronger control conditions upon it than attributionists concede. For the most part, I prefer to allow the case for volitionism to emerge gradually, by way of criticism of attributionism. But I would like to say a word on why I believe that volitionism entails strong epistemic conditions on moral responsibility. It is not sufficient for an agent to be able to cause an alteration in a state of affairs, or even actually to cause such an alteration, for that agent to control that state of affairs. If I do not know either *that* I cause such changes, or *how* I alter the state of affairs, then I do not control it. We causally interact with many, many things without controlling them – the blades of grass we crush and bend as we walk, the air molecules we breathe in and out, and so on. Intuitively, then, if moral responsibility requires control, then it requires that we know what we are doing.

As we have seen, attributionists do not deny that control matters to the assessment of responsibility, but the control conditions they defend are weak, mattering just to the extent to which lack of control can block attribution of attitudes. Consider someone like Robert Harris, the serial killer to whom Watson devotes a long meditation (2004a). Harris may not have had a genuine chance to become a better person; nevertheless the attitudes he expresses in his awful actions are genuinely his. As Scanlon says, of cases like this, if someone commits crimes “because he does not place any value on other people’s lives or interests, what clearer grounds could one have for saying that he is a bad person and behaves wrongly?” (1998: 284).

That people’s actions are typically expressions of their judgments and commitments ought to be common ground. Equally, we ought to agree that some of these attitudes are bad – contemptible, abhorrent, malevolent and so on – and others good. But it neither is, nor ought to be, common ground that we are responsible for everything that can be attributed to us. There is, after all, a natural alternative. We can hold that assessment of attitudes is simply that: the attribution of qualities to the agent that are good or bad, admirable or repugnant. After all, *prima facie* there ought to be conceptual space for such assessment. We ought to be able to say that something is *bad* without saying that it is *blameworthy*.

Attributionists, however, are committed to denying that there is room for such a distinction. Adams and Watson are explicit:

To me it seems strange to say that I do not blame someone though I think poorly of him, believing that his motives are thoroughly selfish. Intuitively, I should have said that thinking poorly of a person in this way *is* a form of unspoken blame (Adams 1983: 21).

In one way, to blame (morally) is to attribute something to a (moral) fault in the agent; therefore, to call conduct shoddy *is* to blame the agent (Watson 2004b: 266).

Presumably because they deny that there is any conceptual room for a distinction between the bad and the blameworthy, at least with regard to the judgment-sensitive attitudes of agents, attributionists devote a good deal of time to defending the claim that such attitudes need not be under the control of agents before we are willing to assess them as good or bad, moral or immoral. As Smith points out, “we do not, in normal circumstances at least, refrain from attributing [such attitudes] to persons or from taking them as legitimate grounds for both moral and nonmoral assessment” (2005: 241). But, from the volitionist point of view, these points are irrelevant to the question at issue. Volitionists agree that we can assess agents upon the basis of their morally relevant attitudes, as the attributionists claim. What they deny is that finding that an agent is morally flawed is necessarily to hold that agent responsible for her flaws; that all negative assessment is blame. After all, volitionists might point out, even attributionists find fault with agents without holding them to be at fault: when these faults are the product of transient mental illness, for instance. It is, therefore, false that there is no conceptual room for a distinction between the bad and the blameworthy, and attributionists owe us an argument for closing the gap between the two.⁴

Given the *prima facie* plausibility of the claim that we ought to be able to distinguish between the *bad* and the *blameworthy*, it seems that the burden of proof falls upon attributionists. They must show that, with regard to judgment-sensitive attitudes at least, there is no such distinction to be made. Attributionists have several arguments designed to shoulder this burden. I shall show that none of them has *any* weight. With regard to each attributionist consideration, volitionists have a rebuttal available, a rebuttal which is sometimes as plausible as the attributionist claim and often more plausible. Indeed, as we shall see, some central attributionist arguments covertly rely upon the very volitionism they seek to counter. Let’s now turn to these attributionist arguments.

⁴ Sarah Buss (personal communication) suggests a variety of attributionism that makes conceptual space for a distinction between the judgment that an attitude is bad and the judgment that it is blameworthy. She argues that if an action is correctly attributable to the agent then she is *accountable* for it, but not necessarily *responsible* for it. Accountability, on Buss’s view, tracks the metaphysics of causation, whereas responsibility is a normative notion. I have no objection to dividing up the conceptual terrain in this manner, but attributionists like Scanlon, Smith and Watson will not be able to accept it. For them, attributability is a normative notion, not a metaphysical one. They owe us an argument showing that there is no conceptual space for a distinction, *within* the normative domain, between actions that are faulty and those for which the agent is at fault.

A. Character

Scanlon makes much of the fact that we can read off what a person is like from their attitudes.⁵ If my action genuinely expresses my attitude, then we can read off the state of my character from it. If my actions consistently express indifference to the suffering of others, for example, then we can conclude that I *am* callous. However, once again it seems *prima facie* plausible to press the distinction between a faulty character and one for which an agent is at fault. Why is there no room for distinguishing between bad and blameworthy characters?

Scanlon might reply that assessments of responsibility must bottom out somewhere, and that the most plausible final court of appeal is in the character of the agent. But why should we think that? Why not say, instead, that responsibility bottoms out with acts of choice? Perhaps Scanlon thinks that that way an infinite regress lies; acts of choice could only ground our responsibility if we were responsible for our ability to choose, and such a demand is in principle unsatisfiable. On this view, only an *ens causa sui* could be responsible. But it certainly does not follow, from the fact that we cannot have chosen to possess the capacity to make choices, that our responsibility does not bottom out in our acts of choice. Suppose, with Wolf (1990), that responsibility requires normative competence. Whether or not agents possess that competence is, typically, not up to them. Instead, it is a matter of luck: Good upbringing, native capacity and so on, is required, and none of these is (initially, at least) within the control of the agent. The fact that agents are not responsible for their normative competence does not, however, prevent us from asking whether they are responsible for their actions. On the volitionist view, I may not be responsible for whether I am in the responsibility ballpark, but I am responsible for what I do when I find myself there.

Indeed, it seems to me that the claim that responsibility bottoms out in the character of the agent can be turned back against the attributionist. Though we are typically not responsible for possessing or failing to possess normative competence, there are occasions upon which this is not so: cases in which once-competent agents have deliberately brought it about that they are no longer competent, perhaps temporarily (cases in which, for instance, agents take a drug designed to temporarily bring about loss of control and extensive character change). The volitionist account allows us to give the right responses to these kinds of cases: If the agent possessed the relevant kind of control over their competence, where “relevant control” includes the satisfaction of relatively stringent epistemic conditions, then she is responsible for her competence or her lack thereof. But it is unclear how the attributionist could even begin to make sense of these cases. Since normative

⁵ Adams’s attributionism seems to rest upon a similar foundation: “No matter how he came by them, his evil beliefs are part of who he is, morally, and make him a fitting object of reproach” (1983: 19).

competence is a central aspect of our characters, she seems forced to say that the agent is *always* and *equally* responsible for it; her character is, after all, always an accurate reflection of itself. It therefore accurately reflects our practical identities.

I do not take this last point to be decisive. It is open to attributionists like Scanlon to deny that we can ever sensibly ask whether agents are responsible for their characters, or that, appearances to the contrary, this question concern only the fairness of imposing sanctions. Nevertheless, the fact that the question seems to be a sensible one just as it stands seems to me to nudge the burden of proof further toward the attributionists' side.

It is a further cost to the attributionist account that it gives counterintuitive results in cases in which character or character change is produced by mental illness or brain damage. If mental illness or brain damage results in transient change of character, Scanlon argues, then we cannot attribute the attitudes it produces to the agent. However, if the alteration is permanent and the agent herself does not reject her new attitudes, then we can attribute them to her. She now has a new character, and is responsible for it (1998: 278-9). On this view, the formerly sober and hardworking Phineas Gage became responsible for his dissolute and anti-social ways, once it was clear that the changes produced in him by the explosion that sent a tamping iron through his skull were permanent. The fact that the damage impaired his ability to control his impulses and altered his sense of what kinds of responses were appropriate is irrelevant to Scanlon. But this seems misguided. Though Gage is plausibly taken to be a disagreeable character, he is not therefore a blameworthy person. Cases like this one illustrate the gulf between the bad and the blameworthy, and force upon us the implausibility of denying the distinction.

Attributionists have considered cases like Gage on many occasions, and they remain unmoved by them. I do not suppose, therefore, that the above remarks will alter their views. However, I suspect they have not grasped just how counterintuitive their view is, when it is applied to cases like this. To bring this out, consider a case in which a nefarious but incredibly skilful and sophisticated neuroscientist intervenes in the brain of an agent to bring about character change, for the worse, in him. Suppose, first, that the character change is permanent. In that case, on the attributionist view, he is blameworthy for the actions, omissions and attitudes that stem from his new character. But suppose that the neuroscientist has the means to reverse the character change. In that case, whether her victim is blameworthy at t depends upon whether the neuroscientist reverses the operation at t_1 (where t_1 is soon after t). On this view, the blameworthiness of the agent depends upon factors entirely outside his control – perhaps on such facts as the state of the neuroscientist's fridge, in which she has stored the brain tissue she has removed. Surely an agent's responsibility should not depend upon facts like that?

Cases like that of Gage also bring out the degree to which relevant control has stronger epistemic conditions than Scanlon acknowledges. Suppose (probably truly) that psychopaths are incapable of at least certain kinds of moral knowledge (Blair 1995; Nichols 2002). Specifically, they cannot grasp the following moral fact: “that an action will hurt someone always counts as a (defeasible) reason against it.” This incapacity makes no difference to their responsibility for hurting others, Scanlon claims: “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” (288). We are able to infer that their character is flawed from the fact that they do not value harms in the right way. It is still the case that their actions stem from “a mode of self-governance [that] has ignored or flouted requirements flowing from another person’s standing as someone to whom justification is owed” (271).

As Watson pointed years before *What We Owe to Each Other* appeared, however, this is false: I can only ignore or flout a requirement if I grasp that requirement (2004a: 234).⁶ That is, I must know that there is such a requirement, and, in addition I must grasp its rationale or have some other reason for thinking it worthy of respect. Scanlon is flat wrong to think (for instance) that a psychopath who injures a person challenges their “standing as someone to whom justification is owed.” If psychopaths do not grasp the moral demand implicit in asking for a justification of an action, they cannot challenge it (except in an externalist sense irrelevant to moral responsibility; the same sense in which a tsunami “challenges” our status as moral agents). In general, ignorance of a moral concern excuses someone of responsibility for failing to consider it. Consider an example: Suppose that there is a kind of harm that is objectively morally relevant, but of which we are ignorant. Suppose, for instance, that plants can be harmed, and that this harm is a moral reason against killing or treading on them. In that case, many of us are causally responsible for a great many moral harms. Are we morally responsible for them? Do we flout a moral requirement, and challenge plants’ standing as objects to which some moral consideration is owed? No to all these questions: If we do not grasp the moral requirement, and this ignorance is not culpable, we do nothing blameworthy. There are therefore epistemic conditions on moral responsibility, which are stronger than anything Scanlon and Smith allow for. Indeed, it seems that the epistemic

⁶ The OED gives as its first sense of “ignore” “Not to know, to be ignorant of.” In this sense, we can indeed ignore a requirement of which we are unaware. But, as the OED goes on to note, this sense of “ignore” is obsolete. Unless Scanlon has decided to revive it without letting us know, we ought to understand him as claiming that the morally bad ignore the requirements of morality in the OED’s third sense: “To refuse to take notice of; not to recognize; to disregard intentionally, leave out of account or consideration, shut one’s eyes to.” But of course *this* sense of ignore does require the relevant knowledge; we cannot ignore that of which we are ignorant. Similarly, we can only flout requirements of which we are aware. OED: “To mock, jeer, insult; to express contempt for, either in word or action.”

conditions on *attributability* are stronger than Scanlon (at least) thinks, since we cannot infer ill will from ignorance in the manner he claims.

In arguing that the wicked person “ignores or flouts” the relevant moral standards, Scanlon seems covertly to appeal to the very volitionist theory he seeks to refute. These are things we do deliberately and wilfully, not blindly and ignorantly. Indeed, the covert appeal to volitional considerations seems to me a constant of his and Smith’s attributionism. Attributionists point out that our judgment-sensitive attitudes are in principle within our control. We do not simply find ourselves with them, in the sense in which we find ourselves with our height or skin color; instead, they are the product of what Scanlon calls our “mode of self-governance” (1998: 271). We produce our judgment-sensitive attitudes; we are *active* with regard to them (Smith 2005: 263). In adopting these attitudes, we take a stand on questions of value (Watson 2004b: 271). Thus these attitudes are sensitive to and the product of our exercise of agency. That is why, when these attitudes are bad, we are not merely held to have faulty commitments or values, but are appropriately *blamed* for them.

But either this is a covert appeal to volitionism or it is simply confused. To be sure, I am responsible for my attitudes if I have genuinely been (relevant) active with regard to them; if I have chosen them. But the attributionist claim is that I am responsible for my judgment-sensitive attitudes not because I have – or even can – control them, but because they are *in principle* “under the control of reason” (Scanlon 1998: 272). It need not be the case that I have controlled them, or even that I could control them, for me to be responsible for them (276). I can see good reasons for thinking that actual control matters morally. But I see none for thinking that *in principle* control matters at all. I don’t have a kind of ersatz control over my car if the steering wheel falls off; the fact that cars are *in principle* controllable does not alter my lack of control in that particular circumstance. If control matters, then its absence cancels responsibility (unless of course the agent is responsible *for* her absence of control). If it does not matter, then we ought not to appeal to it. Appealing to *in principle* control is appealing to an unstable and arbitrary halfway house.

Consider an analogy. Smith and Scanlon believe that the moral views of a psychopath (who does not exercise control over his lack of moral competence) are categorically different from his height because the former belong to the class of things that are in principle sensitive to judgments, whereas the latter does not. But suppose we discover that human beings are a rarity in the galaxy: one of the very few intelligent species whose height is not judgment-sensitive. Do we thereby become responsible for our height? Martian attributionists will claim that we are: Since height belongs to the class of things that are in principle judgment-sensitive, *Homo sapiens*’ actual inability to control their height does not alter their responsibility for it. Martian attributionists would, of course, be wrong: Actual control matters, and in principle control in its absence is simply irrelevant. Note, too, that if it turns

out that we can control our height, we are excused responsibility for failing to do so if we do not know *how* to control it: Moral responsibility requires control, and control requires knowledge.

B. *Justification*

Smith points out that it is often appropriate to ask someone to justify his or her judgment-sensitive attitudes. If someone expresses racist or sexist attitudes, we are entitled to challenge them and ask him or her to withdraw them. Since the relevant attitudes are judgment-sensitive, we can expect that by demanding a justification of them we can influence them. It is in this sense, after all, that agents are active with regard to their judgment-sensitive attitudes: These attitudes are responsive to reasons such that when no such reason is forthcoming, they tend to weaken and eventually extinguish. It is because we are appropriately called to *account for* our judgment sensitive attitudes that we are *accountable* for them.

The problem with invoking the appropriateness of criticism and a call for justification, here, is that the criticism and the demand for justification are both *forward-looking*. When I ask you to justify or abandon an attitude I take to be reprehensible, I draw your attention to what I see as a significant flaw in you, one that you (in principle) can take steps to correct. Compare the case in which you have brought about a state of affairs inadvertently. Suppose that it remains in your power to alter it for the better, but either you are unaware that you have brought it about or you do not believe that it is undesirable (or both). I shall not blame you for your action, but I may very well ask you either to justify it or to modify it. Similarly, when I challenge your attitude, I demand that you take steps to justify it, if you can, or modify if you can't. By drawing your attention to the faulty attitudes, however, I bring it about (or at least take an important step toward bringing it about) that the epistemic conditions on moral responsibility are satisfied.⁷ Since these epistemic conditions matter because they are essential for relevant control, in asking that you justify your attitude I take an important step toward bringing it under your control; if your attitude is now under your control as a result, you are henceforth responsible for it. But, first, you are responsible for your attitude because the volitionist, and not the attributionist, conditions upon responsibility are satisfied, and, second, it hardly follows from the fact that you are *now* responsible that you were responsible all along.

⁷ Merely drawing your attention to a flaw in you may be sufficient to satisfy the epistemic conditions, but is often insufficient by itself. Suppose you have had a racist upbringing. In that case, drawing your attention to your racism is not sufficient to satisfy the epistemic conditions. In addition, I have to draw your attention to the very many strong reasons why racism is false and reprehensible.

C. Apology

The appropriateness of the demand for justification does not support the case for attributionism since, as it is invoked by Scanlon and Smith, justification is a forward-looking notion. But Smith, in particular, does invoke a backward-looking notion: apology. If, as she claims, agents can at any time justifiably be required to apologize for their past attitudes, then it seems to follow that they are responsible for those attitudes. Let us now turn to the argument from apology.

The central case of Smith's paper concerns her failure to remember a friend's birthday, and the subsequent obligation that both she and her friend felt she had to apologize for her lapse. Now, apology and forgiveness are indeed commonly markers of responsibility. Agents offer apologies – rather than excuses or justifications – typically when they wish to acknowledge responsibility for a wrongful action they now regret. Thus, pointing to the fact that an apology seemed appropriate is to the point. However, apologies are not always acknowledgements of responsibility.

As well as acknowledging responsibility for wrongdoing, apologies are also, somewhat paradoxically, conventionally used to signal to others the *absence* of responsibility. If I step on your foot on a crowded subway, we shall both believe that an apology is called for. That apology might take two forms. The first, in this context extremely uncommon, form consists in the acknowledgement of responsibility for wrongdoing. "I'm sorry," I might say. "Just for a moment I thought it might be fun to step on someone's foot. Please forgive me." Much more commonly, I might say, "Sorry, I didn't see you there," or "Sorry, I lost my balance" (so much more common are apologies of the second kind than apologies of the first, we can usually just say "sorry," confident that our speech-act will be understood as an apology of the second kind). In the first case I seek your pardon, and you have the right to give or refuse it. In the second, I seek no such thing. Instead, I explain to you that there was no choice on my part, and therefore no malice. In so doing I let you know that I am *not* responsible. In cases like this, critical responses are *not* justified, for there was no responsibility.

Which kind of apology was called for when Smith forgot her friend's birthday? That depends upon whether she was responsible for the lapse. If there was something that Smith might reasonably have been expected to do, the omission of which caused or contributed to causing her forgetfulness, then an apology of the first kind is appropriate. She should acknowledge her responsibility for her lapse.⁸ But if there was nothing that she might

⁸ Some people will respond that there *was* something that Smith might reasonably have been expected to do, failing which she is responsible for her lapse. Smith ought to have *remembered her friend's birthday*. But of course this claim is simply the claim that is in dispute here. In claiming that Smith's omission was blameworthy only if certain volitional conditions were satisfied, I commit myself to a somewhat controversial view on culpable ignorance. But to my mind that view has been given all the defence it needs by Holly Smith (1983).

reasonably have been expected to do, which would have made her recalling her friend's birthday more probable, then an apology of the second kind is the only one called for. Rather than acknowledging her responsibility – for she has none – Smith should reassure her friend. In these circumstances, an apology has two functions. First, it functions to *deny* responsibility. Smith tells her friend that she genuinely forgot; she did not deliberately slight her. Second, the apology functions to block deep attributability. As Smith notes, when we apologize for our forgetfulness, we typically cite how busy we have been. The point of citing these considerations is to claim that circumstances were sufficiently unusual that our lapse *cannot* be taken to be attributable to us in the way that lapses often are. Because we were so busy, under such pressure, so anxious or so ill, our omission was not reflective of our deep commitments.⁹ As Smith says, we apologize “to reassure the other person(s) that we do still care about these things and judge them to be important” (2005: 248).

Thus, the fact that an apology seems called for is not an indication that the agent who apologizes is responsible (nor, indeed, even that the lapse for which she apologizes is deeply attributable to her). Apologies can be acknowledgements of responsibility, but they can also be *denials* of responsibility. Smith ought to have called her friend and apologized, but, so far as we can tell from the details she provides, she should have done so in order to reassure her friend that she still values her and her welfare, not to acknowledge wrongdoing, for there was nothing she did or omitted to do that brought about her lapse and for which she is blameworthy. She ought to apologize, in order to shore up the friendship and out of consideration for her friend's feelings, not because she is responsible for her lapse.

Indeed, the very fact that she apparently felt it appropriate to offer excuses – to say how busy she had been – by itself demonstrates that not even *she* took herself to be acknowledging responsibility. We sincerely apologize for our wrongdoing only when we admit responsibility for it, and we weaken our acknowledgement to the extent to which we offer excuses, explanations or justifications (Govier & Verwoerd 2002). We offer such explanations to deny responsibility, not to take it. Thus, invoking the felt need to apologize does not establish that we are responsible for whatever can be attributed to us; it doesn't even establish that the lapses for which we ordinarily apologize are attributable to us.

⁹ Thus, an apology for a lapse is always a denial of bare attributability. I apologize either because I acknowledge accountability, and not mere attributability, or in order to deny attributability. Notice the implication that if Smith's account of responsibility for attitudes were correct, then her example would fail: on her account, the link between attitudes and responsibility is defeasible, and in this case the pressures she found herself under when she forgot her friend's birthday would serve to defeat attributability and therefore responsibility.

Weak Attributionism

Neither the argument from character, nor from justification nor apology provides support for strong attributionism, as defended by Scanlon and Smith. By ignoring – or flouting – the epistemic conditions upon moral responsibility, and by failing to acknowledge the distinction between the bad and the blameworthy, they develop an account that, whatever other merits it has, simply is not an account of moral responsibility. What of weak attributionism, as defended by Watson? As we have seen, Watson does allow that moral responsibility has epistemic conditions. Has he thereby shown us how to defend attributionism?

I think not. Though Watson admits that responsibility has epistemic conditions, the conditions he suggests are far too weak to do the job, for reasons we have already glimpsed. For Watson, unlike Scanlon, agents are responsible for attitudes, and for the actions and omissions that express them, only if they are capable of relevant moral knowledge. For Watson, someone is rightly held to be responsible for their cruelty only if they have the concept of cruelty (Watson 2004b: 282), and so on for the other areatic concepts. But, as we have already seen, this is far too weak. For me to be guilty of cruelty, I must know not only what cruelty is, but also that the concept applies *on this occasion* (or be culpable for failing to know it). Consider, once more, the case in which plants suffer harm when they are cut. In that case, perhaps (from the standpoint of some future science, which is able to detect plant suffering) we act cruelly when we pluck roses. But since we don't know that plants can suffer harm when we pluck them, we are not responsible for our cruelty. Simply having the concept is entirely insufficient. The epistemic conditions upon moral responsibility, as we have seen, are in very significant part a function of the control conditions upon which volitionism insists; we need to know what we do in order to be able to control it. But in order to have control over what we do, we need to know how our concepts apply to the case at hand.

Earlier, we invoked Watson to criticize Scanlon's claim that the person who acts badly toward others by that fact alone challenges their moral status. But for exactly the same kind of reasons, we can dismiss Watson's own claim that I take responsibility for my ends simply by adopting them. "To stand for something," Watson writes, "is to take a stand, to be ready to stand up for, to defend, to affirm, to answer for" (2004 b: 271). But this is the case, once again, only if certain epistemic conditions are satisfied. Have I taken a stand on cruelty to plants? I have not, for I have no reason to take the possibility seriously. I take a stand for something against the background of a range of alternatives that I know are taken seriously, as live options, by those around me. I assert the superiority of the ends I affirm, over alternatives. We do not take a stand simply by adopting an end.

Conclusion

All the arguments offered by attributionists are, at best, neutral between attributionism and volitional accounts of responsibility. At worst, they covertly rely upon the very volitionism they aim to refute. The conditions on *attributability* are themselves stronger than attributionists allow; we cannot infer ill will where there is lack of relevant control and when relatively stringent epistemic conditions are not satisfied. But the conditions on genuine moral responsibility are stronger still; even when attitudes are rightly attributed to agents, it is a further question whether they are responsible for them. We care about lapses because, as Smith argues, our lapses are often attributable to us inasmuch as they reveal our commitments. But 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 (character, act or omission) 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.

Moral responsibility is just one part of moral assessment. We can appropriately assess agents, acts, omissions and agents – and perhaps even events – but we can do all of these things without attributing responsibility to anyone. Attributability might be a necessary condition of moral responsibility, but in the absence of relevant control it is not sufficient. There is a distinction between the bad and the blameworthy; because it is unable to acknowledge the importance of this distinction, attributionism fails as an account of moral responsibility.¹⁰

Neil Levy
University of Melbourne
Centre for Applied Philosophy and Public Ethics
nllevy@unimelb.edu.au

¹⁰ I would like to thank Andrei Marmor, Angela Smith, Sarah Buss, two anonymous referees for the *Journal of Ethics & Social Philosophy*, and an audience at the Centre for Applied Philosophy and Public Ethics, Australian National University, for extremely helpful comments on this paper.

References

- Adams, R.M. 1983. "Involuntary Sins," *Philosophical Review* 94: 3-31.
- Blair, R.J. 1995. "A Cognitive Developmental Approach to Morality: Investigating the Psychopath," *Cognition* 57: 1-29.
- Govier, T. & Verwoerd, W. 2002. "The Promise and Pitfalls of Apology." *Journal of Social Philosophy* 33: 67-82.
- Greenspan, P.S. 2001. "Genes, Electrotransmitters, and Free Will," in David Wasserman and Robert Wachbroit (eds), *Genetics and Criminal Behavior* (New York: Cambridge University Press), 243-258.
- Nichols, S. 2002. "How Psychopaths Threaten Moral Rationalism, or Is it Irrational to Be Amoral?" *The Monist* 85: 285-304.
- Scanlon, T. 1998. *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
- Smith, H. 1983. "Culpable Ignorance." *The Philosophical Review* 92: 543-571.
- Smith, A.M. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life" *Ethics* 115: 236-271.
- Strawson, P. 1962. "Freedom and resentment." *Proceedings of the British Academy* 48:1-25.
- Watson, G. 2004a. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme," in *Agency and Answerability* (Oxford: Clarendon Press), pp. 219-259.
- Watson, G. 2004b. "Two Faces of Responsibility," in *Agency and Answerability* (Oxford: Clarendon Press, 2004), pp. 260-288.
- Williams, B. 1981. "Moral Luck," in *Moral Luck: Philosophical Papers 1973-1980* (Cambridge: Cambridge University Press), pp. 20-39.
- Wolf, S. 1990. *Freedom Within Reason*. New York: Oxford University Press.