

General Aspects of Law **GALA**

**GOLDBERG ROOM
(297 SIMON HALL)**

**THURSDAY
APRIL 8, 2010 - 12:10 P.M.**

*“Being Responsible, Taking Responsibility,
and Penumbral Agency”*

By



DAVID ENOCH

**DAVID ENOCH, ASSOCIATE PROFESSOR OF PHILOSOPHY
AND JACOB I. BERMAN ASSOCIATE PROFESSOR OF LAW
THE HEBREW UNIVERSITY OF JERUSALEM***

NOTE

Discussion will proceed without a presentation of the paper by the author. Papers are distributed to those on our mailing list or electronically when available at the GALA website:

< <http://www.law.berkeley.edu/1151.htm> >.

Others may obtain copies from Ms. Amatullah Alaji-Sabrie at 510.642.3627 or asabrie@law.berkeley.edu. Copies can be made available in an alternate format upon request.

GALA events are wheelchair accessible.

For any disability-related accommodations advance notice is requested.

*During the academic year 2009-2010, Harrington Fellow, The University of Texas

(Forthcoming – after revisions that are yet to be made – in Ulrike Heuer and Gerald Lang (eds.) *Luck, Value and Commitment: Themes from the Ethics of Bernard Williams*, OUP 2011.)

Being Responsible, Taking Responsibility, and Penumbral Agency

David Enoch*

1. An Autobiographical Prelude about Bernard Williams

For those of us who are hardly ever convinced by Bernard Williams' conclusions, his greatness is evident not so much in the rigor or precision of his arguments, as in the force of his insights. For many years now, I have been convinced that Williams' discussion of moral luck¹ is seriously flawed, and that his arguments are unclear and – to the extent that I nevertheless understand them – unpersuasive. And yet unlike with other texts whose arguments I find unclear and unpersuasive, throughout these years I have not been able to set these texts aside in my mind. There is, for me, a lingering worry that the flaws of his arguments notwithstanding, Williams was on to something, and furthermore something deep, revealing, and important.

This is a feeling I often get in reading and thinking about Williams' texts. In the context of the moral luck debate, I think this feeling is focused (again, for me) on Williams' introduction into moral philosophy of *agent-regret*: Both a driver and a bystander can equally regret the loss of a life of a pedestrian that the driver hit, but only the driver can, is likely to, and probably should feel that extra kind of bad feeling, perhaps because her agency was involved in bringing about the pedestrian's death. Of course, if the driver believes herself to be somehow *guilty* of the death (perhaps she was driving terribly negligently), then guilt may be the thing she can and the bystander typically cannot feel. But

* For helpful comments and discussions, I'd like to thank Sarah Broadie, Peter Cane, Alon Harel, Ulrike Heuer, Bob Kane, Joseph Raz, , and the students at Joseph Raz's Legal Theory seminar (Columbia, fall 2009).

¹ See Williams 1976 and 1993.

crucially for Williams, there seems to be some kind of feeling that is only available to the driver even if she believes she was *not at all* driving negligently.

It is not completely clear what to make of this phenomenon. I am sure that interesting psychological questions can be (and have been) asked about it, for instance. But for me, the main worry is philosophical, and it comes from the *normative* status of agent regret. For agent regret seems not only common, but also called-for. As Williams writes:

What degree of such feeling is appropriate, and what attempts at reparative action or substitutes for it, are questions for particular cases, and that there is room in the area for irrational and self-punitive excess, no one is likely to deny. But equally it would be a kind of insanity never to experience sentiments of this kind towards anyone, and it would be an insane concept of rationality which insisted that a rational person never would. (29)

Williams seems to think that important conclusions follow from this observation, and here again I think that his arguments lack in precision and rigor what they have in insight. But a worry remains – if, as I believe, there is no moral luck, isn't it surprising that agent-regret should be called-for (or even just rationally permissible), that there is something morally amiss with the non-negligent driver who kills a pedestrian, and feels nothing except for the sorrow the bystander also feels?

The claims I am about to make are broader in scope, and they apply not only to the moral luck debate. And my arguments are going to take me far from Williams' texts. But for me, thinking about these things starts from the need to deal with the lingering worry that Williams – wrong though he probably was – was here too really on to something.

2. Three (Arguably) Related Kinds of Cases

I'm afraid it is you. You are the driver from Williams' example. You hit a pedestrian causing him serious harm. But the accident was not your fault (perhaps it was the pedestrian's, or someone else's, or no one's fault at all). And you know as much – we can stipulate that you're as certain of this as of anything. Now, I do not believe in moral luck, and in particular I do not believe that the mere causal story here – the fact that your agency was a part of the causal chain of events that led to the injury – matters when it comes to your moral record. You are not, it seems to me, morally responsible for the injury. Suppose – for the sake of argument – that my reasons for so believing² are at least prima facie compelling. Still, Williams was on to something. While we should all feel bad for the fate of the injured pedestrian, you, it seems, should feel that extra bit of agent-regret. If you do not, we will tend to judge you unfavorably, and there will be a distinctly moral flavor to this judging (“I understand it wasn't her fault, but still, what kind of person just walks away from such an event and proceeds with her life as if it was nothing to her?”). At the very least, we do not think that this kind of behavior is the uniquely rational response to the situation, we think that it is at the very least rationally and morally permissible to feel agent-regret.

Your teen-aged son commits a crime, causing harm to person and property. You are not, let us suppose, directly responsible for the crime in any straightforward way – it's not as if you put him up to it, or even drove him to this kind of thing by your poor parenting. Parenting too, after all, is a percentage game, and this time you lost. Now, I am an individualist about moral responsibility – I believe that you cannot be morally responsible for something simply in virtue of being in a certain relationship with someone else who is morally responsible for that thing; in morality – perhaps unlike the law – real vicarious liability is impossible. Suppose – for the sake of argument – that my reasons for this individualism are at least prima facie compelling. Still, we would judge unfavorably a parent who neglects to – in some sense – take responsibility for her teen-aged son's behavior, perhaps, for instance, by apologizing for him, or some such. If facing hard questions, you settle for noting (correctly) that the

² Some of which are in Enoch and Marmor (2007).

relevant action was not yours, there seems to be something amiss – I would say, morally amiss – with your so doing. It's not that what you're saying is wrong: the action, after all, *was* not yours, and you *were* not responsible for it. It's that your *saying* so is somehow problematic. At the very least, we do not think that this is the only rational response here. There is something to be said for your (in some sense) taking responsibility for your teen-aged son's action. Or so, at least, it seems to me.

Your country has recently been engaging in some morally problematic projects and actions, perhaps more so than the extent to which this is true of just about any country at any given time. You are not – let us suppose – directly responsible for any of this. You have not yourself taken active part in the problematic activities, you haven't even voted for the people perpetuating it. You may have even gone to a couple of demonstrations protesting against it. Once again, I am an individualist when it comes to moral responsibility: I certainly do not believe that merely in virtue of being a member in a collective, or a citizen of a country, your moral record is automatically stained when the collective's or the country's is³. Suppose – for the sake of argument – that my reasons for this individualism are at least *prima facie* compelling. Still, if the opportunity arises when moral charges are brought against your country, there would be something wrong, it seems to me, if all you did was to (correctly) point out that these actions are not yours. To repeat, it's not that this would be false – these are indeed not your actions, and you're not responsible for them. Rather, it's that your response should at the very least be more complicated than this. At least if some other background conditions are in place, you should, I think, *take* responsibility for your country's actions, and, say, attempt to justify them, or excuse them, or apologize for them, or some such. If you refuse to do any of this, and instead just settle for noting that the relevant action is not yours, there is, I think, something morally amiss with your response. At the

³ Notice that the issue relevant here is not when (if ever) a collective is responsible for anything. The issue is, rather: *Given* that the collective is responsible, what does it take for a (personally uninvolved) individual to be responsible partly in virtue of being a member of a collective that is responsible for the relevant thing. In the terms sometimes used in this literature (see, for instance, Smiley (2005, section 4)), the issue is that of the *distribution* of collective responsibility, not that of its *existence* or emergence.

very least, this is not the *only* acceptable response. There's something to be said for taking responsibility for one's country's actions.

Of course, nothing here is obvious or uncontroversial. For one thing, you may not find at all convincing some of my intuitive starting points, the judgments that in these three cases failing to (in a sense yet to be made more precise) take responsibility is morally problematic⁴. Or you may agree with me on these points, but differ on the suggestion that these three kinds of cases are instances of an interestingly uniform normative phenomenon. True, I have done what I can to present them as at the very least closely analogous, but it's not as if I've presented an argument to the claim that they are so closely related (hence the "arguably" in the title of this section). And even if you do agree with me on the essentials, there is still a lot of work to be done: We may ask, for instance, what is included in the requirement (if it is a requirement) to take responsibility for the consequences of one's actions? And at which ages (or which developmental stage) of the child does the parent's duty to take responsibility for his son's actions phase out? And what conditions must a country satisfy for it to have the kind of relation with its citizens that calls for their taking responsibility for its actions?

So I do not want to pretend that anything here is simple. And I'm afraid I will not in this paper have much to say by way of the rich and complicated details needed for replies to the last three questions and to many others. But I do want to insist that there is something similar to the three cases described, and in what follows I want to offer an initial and general analysis of the phenomenon present here. Intuitively, in all these cases the thing for which responsibility is in question – the consequences of your actions, the actions of your son, or of your country – is not within the core of your agency. But it is not too far from it either. As I will put things, we are here in the area of your penumbral agency. This is

⁴ Strictly speaking, all that I need from these examples is that there is an important difference here between the one taking responsibility and the one failing to do so, so that the former behaves in a morally better way. This may be because the former acts in accordance with the duty to take responsibility and the latter doesn't. But – as was pointed out to me by Eve Heafey – this may also be because the former behaves in a supererogatory way and the latter isn't. In the text I assume – as seems likely to me – that there is here a duty to take responsibility. But everything I have to say survives pretty much unchanged if we go for the supererogation reading instead.

not just the borderline area in between things that determinately are and things that determinately are not manifestations of your agency. “Agency”, I’m sure, is vague much like any other natural-language term. But what is going on here is not just vagueness. Rather, penumbral agency is a normatively different *kind* of phenomenon. The normative implications of something being a manifestation of your penumbral agency are different in kind from both cases of agency and cases of non-agency. And there may be cases of borderline agency that are not cases of penumbral agency.

In such cases of penumbral agency we have independent reasons – having to do with the control condition on moral responsibility, mostly – not to assign full-blooded moral responsibility for the relevant action or event. This will, in most parts of this paper, serve as an un-argued-for premise. But this will not entail that there’s nothing responsibility-related in the vicinity here. Rather, in all these cases you are plausibly morally required to *take* responsibility, where this taking makes a difference to the responsibility-facts.

Because of the necessarily partial nature of the discussion to come, I take myself here to establish more a possibility-result than an outright conclusion. Thus, I will not defend the claim that in these cases you are not responsible but rather you are under a moral duty to take responsibility. Rather, I will defend the claim that this is a coherent conjunction, and furthermore a potentially productive one, one that seems to capture something dear to our heart in the phenomenology of responsibility. This is why even if you don’t find some of the judgments above initially plausible, you should continue reading: even if I am wrong about one or two of the three cases, still I may be right about the other(s), or about other cases not mentioned here. And anyway, the conjunction I am after (no responsibility, but duty to take responsibility) may be coherent regardless of our ability to come up with plausible examples for its realization (though if we cannot come up with such examples, of course, it will be much less interesting). And if you just cannot see the three cases as instances of the same (or a similar) phenomenon, I suggest that you wait to see if this changes after my account has been presented. By then, I hope to convince

you that there *is* a coherent, normatively interesting, and fairly uniform phenomenon common in all these cases. Sometimes the uniformity of explananda becomes clearer only after their common explanans is in reasonably clear view.

3. Taking Responsibility: The Intuitive Thought

The intuitive thought that I will attempt to elaborate on in the rest of this paper, then – and which, to an extent, I borrow from Susan Wolf – is the following⁵. Some things are within the core scope of your agency. Your actions, perhaps, or your intentions, are such that you are responsible for them.

Furthermore, although it is often up to you whether to perform an action, or form an intention, and although their being up to you may be importantly related to your being responsible for them, your being responsible for them is not up to you. Once you perform the action, or form the intention, and assuming the other necessary conditions are in place, you are *already* responsible. Your being responsible does not depend on any act of will (except, that is, insofar as the action or intention or whatever it is you are responsible for so depends).

Other things lie outside the scope of your agency. The motion of the planets, say, or the (natural) length of your nose, or the actions of some far away people to whom you stand in no special relation are all things for which you are not responsible. If the topic comes up, if someone treats you as if you were in some way responsible for these things, then saying that you are not, that these are not actions of yours is *precisely* the right response on your behalf. In these cases, unlike the three cases I

⁵ This intuitive thought is the guiding one in Wolf's "The Moral of Moral Luck" (2001, throughout, but see especially 121). Though there are important differences between the view I present in this paper and Wolf's, and though many of the details differ (some of these differences are highlighted below), still the central intuition – certainly, the central phenomenological analysis – is already present in Wolf's paper, for which I am indebted. Let me highlight here one difference: Wolf is almost entirely concerned in that paper with the phenomenon of moral luck. When she discusses other examples (quite similar to the second and third examples I started with) she does this anecdotally and tentatively (2001, 121). For me, though, the relevance of the other examples is crucial, as the suggestion that there is here a unified phenomenon is central in supporting the analysis I will eventually offer. To an extent, something close to this guiding thought is also present in Adams (1985), though Adams would not, I think, accept the voluntarist part of my account, according to which in the relevant cases your taking responsibility *makes it the case* that you indeed are.

started with, nothing else is called for – you are just not responsible, and that is the end of the matter. Furthermore, there’s nothing you can do in order to change this. Even if you want to become responsible for them, you cannot – it is not, one may say, within your power to take responsibility for them. (This statement should be qualified. Presumably there are senses of “responsibility” in which one can take responsibility even for these things. But these are senses which will not interest me. I discuss different senses of “responsibility” drawing some necessary distinctions in section 5.)

But there is a third group of cases. These are cases of things that lie in the penumbra of your agency. In such cases – arguably exemplified by the three examples I started with – you are not responsible for the relevant thing (the uncontrolled consequences of your actions, your child’s actions, your country’s actions). But in these cases, an act of will can make all the difference. In these cases, you can change the situation with regard to your being responsible. You can, that is, *take* responsibility, and thereby *become* responsible⁶.

Furthermore, taking responsibility is a kind of an action⁷, and so it may have one of the moral statuses that actions often have. In particular, it can be the case that you *ought* to take responsibility. And sometimes, this is precisely the case. Indeed, I think it is the case in the examples I started with. You are not, I think, responsible for the injury to the pedestrian. But you ought to take responsibility. And once you do, you will be responsible, partly in virtue of your act of will, your having taken responsibility. Only partly, because the act of will, though necessary, is not sufficient – it is also necessary for the case to antecedently lie in the penumbra of your agency: If it does not, you cannot bring about a change in your responsibility simply by attempting to take responsibility. But noticing the possibility of a duty to

⁶ Here are two other important differences between me and Wolf (2001): First, I think it is a central part of the phenomenon we both seem to be after that it involves an act of will. Wolf nowhere talks about the taking or responsibility as an act of will. Second, responsibility is not as central to Wolf’s project as it is to mine, and so she does not give any details about the relevant kind of responsibility – details of the kind I give below.

⁷ This is one major way in which my use of “taking responsibility” differs dramatically from Fischer and Ravizza’s. See, for instance, Fischer and Ravizza (2000, 443) and the references there. Note that despite explicitly saying that taking responsibility is, as they understand it, not an action (but rather as having certain dispositional beliefs and having appropriately acquired them), they still think about it as a way of making an action one’s own. But the metaphors here are somewhat misleading, because “making one’s own” very much suggests an action.

take responsibility, I want to argue, opens up some room in logical space for another way of accommodating our negative judgment of the driver who just proceeds as if not much has happened. Her moral flaw is not constituted by her being responsible for the pedestrian's injury and failing to respond appropriately (she is, ex hypothesis, not responsible). Rather, it lies in her violation of her moral duty to take responsibility.

Similarly, I would say, for the other two examples. You are not, I think, responsible for your teen-aged son's crime. But you are – in certain circumstances, at least – under a moral duty to take responsibility for it. If you do, you are then responsible for it, partly in virtue of having taken responsibility for it. If you do not, you are not responsible (you did not, after all, take responsibility for it, and his action is not within your core agency). So you are off the hook – off *this* hook, that is. But you are still acting in a morally problematic way, because you are violating your moral duty to take responsibility for your son's actions. And this is what explains our negative evaluation of the parent who in the face of such circumstances proceeds to just (rightly) note that the teen-ager's actions are not his own.

And you see how it goes: You are not responsible for your country's actions. But you are – in certain circumstances – under a moral duty to take responsibility for it. If you do, you are then responsible for it, partly in virtue of having taken responsibility for it. If you do not, you are not responsible for your country's action, but you are in violation of your moral duty to take responsibility for it. And you may very well – presumably depending on some other considerations – be responsible for *that*. And this is what explains our negative evaluation of the citizen who proceeds in such a case simply to (rightly) note that her country's actions are not her own.

This kind of thing sounds to me like the right thing to say – I think that the last three paragraphs capture rather nicely the phenomenology of these cases. In all of them, it seems to me that being responsible is not something we find ourselves with, but rather something that we do, there is something distinctively active about becoming responsible in these cases. This, then, is what it seems to

me we *want* to say about the three cases, at least in outline. But of course, in philosophy as anywhere else, you can't always get what you want. In order to see whether we *can* responsibly say such things, we need many more details. In particular, we need an account of what taking responsibility consists in, and an account of what the responsibility thus taken is. These are the tasks of the following two sections.

4. The Taking

Sometimes, when we use the locution "taking responsibility" all we mean by it is something epistemic in nature. If I say "It's time that you start taking responsibility for your actions!" what I typically have in mind is not something about penumbral agency at all. What I typically have in mind is that it's time that you *acknowledge* or *recognize* the relevant responsibility-facts, namely, that you are responsible for your actions (in general, or perhaps for some particular actions). When "taking responsibility" is thus used, the responsibility facts in no way depend on the taking. Rather, they are there all along, and the taking is thought of as the appropriate response to the relevant responsibility-facts. In this sense, then, taking responsibility is analogous to acknowledging a duty that is already there. And there is, of course, nothing wrong with using "taking responsibility" in this way. But if you share my intuition that sometimes we can take responsibility even in cases in which (as we know) we are not responsible independently of and prior to the taking, then you already agree that this epistemic sense of "taking responsibility" is not the only relevant one.

Another way of using this phrase is the way I want to focus on. This is the way in which taking responsibility is analogous not to acknowledging a duty that is already there independently of the acknowledging, but rather to *undertaking* a duty (or obligation, or commitment). In this sense, when I take responsibility for something – and when the other necessary conditions are in place – I thereby *become* responsible. Having taken responsibility (in this sense) I *am* responsible, and this is so partly in

virtue of my having taken responsibility. Think again about the example of undertaking a commitment, say by making a promise. When I promise to meet you for lunch, typically the promise does not amount to acknowledging a duty (to meet you for lunch) that was there anyway, independently of the promise. There *was* no such duty independently of and prior to the making of the promise⁸. Rather, the making of the promise created the obligation⁹, and now I am under a moral duty to meet you for lunch partly in virtue of having made that promise.

I am going to make rather heavy use of the promise analogy, so let me stress already here that it is *only* an analogy. As is often noted (and as will be further discussed in the next section), talk of responsibility is multiply ambiguous. Sometimes our talk of responsibility is forward-looking, and then it seems to just be equivalent to talk of duties, or perhaps some kind of duties (this is the kind Baier (1984, 104-5) calls “task-responsibility”). This seems to be the sense of “responsibility” in play when we say that it is your responsibility to bring a salad to tomorrow’s picnic. And there is a corresponding sense of “taking responsibility” as well, as when we ask “Who takes the responsibility for bringing the salad for the picnic?”. If you proceed to take responsibility for bringing the salad, what you are doing seems to be to undertake a commitment to bring the salad (perhaps it’s a promise, or perhaps some close relative thereof). And again, there is nothing wrong in this way of using words. But it is important to note that this too is not the sense of “taking responsibility” I am after, because this is not the sense of “responsibility” I am after. I am after the (or a) *backward-looking* sense of responsibility. I will have more to say on this in the next section, but for now it is just important to remember that it is not the forward-sense of “responsibility” in the taking of which I am interested. This is why the promise analogy is merely an analogy, and should be treated with some suspicion (a point to which I return shortly).

⁸ Depending on your theory of promises, and also, I would say, on your theory of reason-giving in general, there may have been *a* duty there independently of the making of the promise, namely, my duty to keep my promises, and so perhaps also my duty to-meet-you-for-lunch-if-I-so-promise. See my “Giving Practical Reasons” (manuscript). But this doesn’t matter for our purposes: All that matters here is that the duty *to meet you for lunch* did not exist independently of and prior to the making of the promise.

⁹ Throughout the paper I am using “duty” and “obligation” interchangeably.

Having warned you about the analogy, let me exploit it a little further. When making a promise (say, to meet you for lunch), typically – and as already noted – the normative situation will be as follows: Independently of and prior to the promise, I am not morally required to meet you for lunch. The making of the promise, though, changes this, and now I am morally required to meet you for lunch, partly in virtue of having promised to do so. But let me note a further interesting feature of promises – one that is not, I think, often noticed. The making of a promise is itself an action, and so we may wonder about its moral status. In particular, I may be morally required to make a promise. And this is consistent with my not being morally required to perform the action that I am morally required to promise to perform. Here's a plausible (though somewhat melodramatic; I apologize) example: A single-parent friend of mine is diagnosed with severe illness. He is extremely anxious about who will take care of his daughter if he dies. Suppose that we are not close enough for it to be morally required that I take care of her. Indeed, suppose that if my friend dies without anyone having committed to taking care of his daughter, still one or two closer friends will step up to the plate. But suppose further that none of them will promise my friend now that they will. By promising, I can significantly alleviate my friend's anxiety. It seems to me that there are ways of filling in the details of the example such that I will be morally required to promise my friend to take care of his daughter if he dies. This, even though I am not morally required – independently of and prior to the promise – to take care of her if he dies. Of course, if I make the promise and the friend later dies, I am then morally required to take care of his daughter, partly in virtue of having made the promise. If I neither make the promise nor take care of the child, I am not in violation of my moral duty to take care of the child (I don't have *that* duty). But I *am* in violation of my moral duty to *promise* to take care of her – a duty I can be in violation of even if I do proceed to take care of the child.

Here too I do not want to pretend that things are simple or obvious (for one thing, as it stands the example depends on a moral duty to keep promises to people even after they no longer exist, and I

have no idea whether we have such a duty). But the crucial thing for our purposes here is not whether the example is convincing morally, but whether it is coherent, and so manages to capture a coherent normative structure. And it seems to me clear that it is: It is one thing to have a duty to ϕ , quite another to have a duty to promise to ϕ . The normative considerations that apply to one of these need not apply to the other, certainly not as a matter of necessity. And there is no general formal reason to think that you cannot have a duty to promise to ϕ without also having a duty to ϕ (of course, if you act in accordance with your duty to promise to ϕ , you *then* have a duty to ϕ ; this, after all, is precisely my point). Even if you reject the melodramatic example above on substantive grounds, then, still it seems to me you should acknowledge the formal coherence of the normative structure I am after.

It is precisely the analogy with this normative structure that I want to exploit. In the cases of taking responsibility of the kind I'm interested in, I want to say, there is no responsibility (independently of and prior to the taking); there is a duty to take responsibility; if you take responsibility, say, for your child's actions, you thereby become responsible, partly in virtue of your having taken responsibility; and if you do not take responsibility, you are not responsible for the child's actions, but you are in violation of your duty to take responsibility, and perhaps also responsible for *that*. And now that we know that there is no formal problem with this normative structure when it comes to promises, we should not be too worried about this structure in the case of responsibility either. (Though this is too quick – responsibility may not be like promises in the relevant ways. I return to this worry shortly.)

Let me use the promise analogy also to make another point. Earlier on I insisted that some responsibility-takings are just not within your power. You cannot take responsibility for the motions of the planets, for instance, or for the actions of people you stand in no interesting relation to. This is why when you take responsibility, you are then responsible only partly in virtue of your act of will – some other conditions must be in place for the attempt to take responsibility to succeed. And this phenomenon too has a close analogue in the case of promises. After all, it's not as if any (even sincere)

utterance of “I hereby promise to ϕ ” is guaranteed to succeed in generating a duty. If we replace “kill your first-born” for “ ϕ ”, for instance, or if the utterer is utterly insane, then even if there is here a genuine attempt to generate a duty to ϕ , it will not succeed. In the case of promises too, then, the duty is only partly grounded in the relevant act of will – here too there are other necessary conditions.

What goes on, then, in the cases of taking responsibility of the kind I am interested in – the kind that is arguably involved in the three examples I started with – is analogous in these ways to what goes on when we make promises. But it’s time to revisit the worry that I am making much too much of this analogy. After all, what *argument* have I given, what reason have I cited to believe that taking responsibility *is* like promises in the relevant way? This is a fair challenge, of course. But let me remind you of my rather limited ambitions in this paper. I am trying to establish a possibility, room in logical space, and I am hoping that when we consider the three examples I started with this further logical possibility in mind, it will seem plausible to us – true to the phenomenology of the examples – that they occupy that room in logical space. I do not pretend here to offer a conclusive argument for any stronger conclusion. So noting that the normative structure I am after is one that is already plausibly instantiated in the case of promises is a good first step on the way to establishing the possibility result. And if we then think about the examples and the room just opened in logical space seems very relevant to them, then this concludes (this part of) my modest project. I am not sure that more can be done, but I am quite sure that I am not trying to do more here.

But isn’t there an important *dis*-analogy between the two cases? When you make a promise, this changes your relevant normative status (it places you under a duty) only from this point in time onward. Doesn’t taking responsibility bring about a retroactive change, in that if (for instance) today you take responsibility for your country’s actions, you are (if I am correct about taking responsibility) now responsible for your country’s actions *of yesterday* as well? How is that even possible? Isn’t this some weird kind of backward-causation, or something too close to it for comfort? At the very least, isn’t this

retroactivity sufficient to break the superficial analogy I have been relying on between taking responsibility and making a promise?

The answer to all of these questions is “no”. True, the action (or whatever) for which I take responsibility at t_1 may have occurred earlier, at t_0 . But the brought-about normative change is still prospective: It is only from t_1 on that I am responsible for what happened at t_0 . And here too there is a close analogue in the promise-family. If you do not know who won the game at t_0 , but want to be able to rely on knowledge of who the winner was, and if I then tell you at t_1 “The Longhorns won. I guarantee it.”, it then becomes a moral failure of mine – *from t_1 onwards* – if the Longhorns lost at t_0 . So nothing mysterious or unique is involved in the possibly backwards-looking nature of taking responsibility, and no disanalogy with promises has been presented. Now, there is still *some* retroactive feel to taking responsibility (in my sense). Roughly speaking, it is that by taking responsibility one retroactively incorporates a possibly past event from the scope of one’s penumbral agency into one’s agency, thus making it (in a sense) one’s own. But I don’t see any reason to think of this kind of retroactivity – about which I will say more in section 5 – as a serious problem.

You may still be unconvinced. You may think, for instance, that while whether we are under a duty may be up to us (and so it is sometimes within our power to make promises), it is just never up to us whether we are responsible for something. We either are, or we are not, and either way there is nothing we can do to change things in this regard. In order to deal with this worry, it will be necessary to distinguish the senses of “responsibility” for which it is well-taken from those for which it is not. I do this, and I argue that the case at hand is an instance of the latter, in the next section.

Before doing that, though, I want to spend some time on the technicalities here – how it is that we take, or fail to take, responsibility.

The first thing to note here is that while we can certainly take responsibility explicitly (“I hereby take responsibility for my son’s actions.”), we much more often do so implicitly. And the many ways in

which we can do so implicitly differ widely from one context to another. Here too the analogy with promising is telling. There is hardly any way in which we cannot – in *some* context – make a promise. Sometime, one’s intonation can make the difference – for instance, sometime the tone, and maybe also the body language, with which a simple “don’t worry” is uttered suffices for it to qualify as a promise, as genuinely duty-creating as the most solemn “I hereby promise”. Indeed, sometimes there is no need for any text at all, and the promise is just implied by a certain behavior. And furthermore – and this is going to be important when I return to taking responsibility – sometimes a promise can be implied by the initiation of a chain of actions that constitute keeping that promise. Thus, I can promise to take care of your college tuition, and then when the day comes start making payments. The first payment will then be the first part of keeping the promise. But we can easily imagine circumstances where there is no promise preceding the first payment, and the first payment itself embodies the (implicit) promise to take care of the other payments as well, that is to take care of your college tuition as a whole. In such a case, the first step in keeping a promise can also embody the making of that promise.

Getting back to taking responsibility, then: It too, as I already said, can be done implicitly. And just as with promises, there is hardly any limit as to what can qualify as an implicit way of taking responsibility – in the appropriate circumstances, a meaningful look, a body gesture, a text seemingly doing some other thing can amount to a taking of responsibility. And here too, sometimes one can implicitly take responsibility by performing the first part of what it would take to live up to the responsibility just taken. In the next section I am going to argue that central ways of living up to the responsibility taken include things like offering a justification for the relevant action, or an excuse for it, or an apology. What I want to note here, though, is that doing such things can also constitute implicit ways of taking responsibility, as well as the first step towards living up to the responsibility thus taken. So, for instance, by offering an apology for your child’s behavior you can at once both take responsibility

for it, become responsible for it, and then go a long way towards acting in the way one who is so responsible should.

This complicates things, of course, because it makes it harder to distinguish between the hypothesis that one is – notwithstanding what I took as a premise earlier on – quite simply responsible for one’s country actions (say), and the hypothesis that one is not, but that one is under a moral duty to take responsibility. Once we see that on the latter hypothesis the very same act can serve both as the taking of the responsibility and as partly discharging it, it may be hard to see what the difference between the two is. Does it matter, after all, if we go for a theory that judges an act of will (taking responsibility) redundant, unnecessary for responsibility, or for a theory that considers such an act of will necessary for responsibility in these cases, but is willing to read this act of will into just about anything, including the very action that (both theories agree) is a major part of what is required of someone who is already responsible? But differences between the two types of theory remain. For the two hypotheses differ regarding what it is that is wrong about the relevant agent’s morally faulty behavior. On the first hypothesis, what is wrong in the behavior is that it fails to respond in the appropriate way to the will-independent responsibility facts. On the hypothesis suggested here, though, the flaw is the failure to live up to the duty to take responsibility. Consider, for instance, the injuring driver example again. And assume for now that one of the things a responsible driver should do in such cases (once she is responsible for the injury, that is) is to repeatedly visit the injured pedestrian in the hospital. If you fail to visit the pedestrian you (non-negligently) hit, what is the flaw in your behavior, then? On the hypothesis that you are quite simply responsible for his injury, what is wrong is that you do not respond in the appropriate way to your being responsible for the injury. On the theory that I’m suggesting, though, what is wrong with your behavior is that you fail to take responsibility, and then live up to being responsible. And quite possibly, the natural way to take responsibility here would be to make the first trip to the hospital. On the suggested theory, then, there is an interesting difference

between the moral flaw involved in not coming at all, and the moral flaw involved in coming for the first time, but never again. Not visiting at all amounts (on our assumptions, and assuming that no other surprising features are present, like a detailed “I hereby take responsibility and apologize” letter) to a violation of the moral duty to take responsibility. Coming once may very well amount (depending on the circumstances) to taking responsibility. From that point on, not coming again amounts to failing to respond appropriately to the fact that one is (now) responsible for the injury in the relevant sense. Intuitively, this sounds exactly right to me – there *is* a difference between the two flaws involved in these two kinds of behavior. The first theory – according to which you just are responsible for the injury, regardless of any act of taking responsibility – sees no such difference. So this is an advantage of the theory suggested here.

Here is another way of seeing a difference between the two understandings of responsibility in this kind of case, a difference that counts in favor of the suggested account of taking responsibility. Think of two citizens, then, Proud and Detached, who both refuse to take responsibility – in either my sense, of refusing to assume the responsibility, or in the competing sense, of just refusing to acknowledge that they are already responsible and respond accordingly; and suppose that these two citizens are alike in all relevant respects (their countries are equally democratic, the wrongs committed by the country equally grievous, both citizens enjoy equally the benefits that their countries offer them, and so on), except one: Proud takes pride in the achievements (moral and otherwise) of his country. Detached, on the other hand, feels detached – she can be happy, of course, for some of her compatriots when they do well, even when they succeed in what may be considered national projects. But the national success itself is nothing to her, as it were. It seems to me clear that when we judge Proud’s and Detached’s refusal to take (or acknowledge) responsibility, we want to distinguish between them. There is something *much* more disturbing about Proud’s attitude here than about Detached’s. Of course, on our assumptions – that this too is a case where on the suggested theory Detached is under a moral duty

to take responsibility – Detached’s behavior is not morally ideal. She too violated a moral requirement that applies to her. But something worse – or at the very least something significantly *different* – seems to be going on in Proud’s case. Having taken pride in his country’s achievements, shouldn’t he now accept the responsibility for its wrongdoings as well? If we understand what is going on here in terms of the duty to take responsibility, and if we understand responsibility here as partly due to an act of will (that is, the taking), we have a ready explanation of the phenomenon here. By taking pride in his country’s achievements, Proud has already (implicitly) taken responsibility for his country’s actions more generally, he has already incorporated this activity from the penumbra of his agency into his agency. So when he now says that he is not responsible for his country’s wrongdoing, what he says is strictly speaking false: Having taken responsibility, he *is* now responsible. So by his denial he in effect evades the responsibility that is already there, he fails to respond to it correctly. Detached, on the other hand, has not taken responsibility, and so she is not responsible. True, she fails to take responsibility, as we are assuming she should. But she does not evade a responsibility that is already there. And there may be a difference in the degree of severity (in general, or more plausibly, in the specific circumstances) between the two wrongs. At least, on the suggested understanding of taking responsibility Proud’s and Detached’s behavior are different in a normatively significant way. And this is precisely the result we were looking for¹⁰. On the competing suggestion currently on the table, however, it is hard to come up with a similarly clean explanation of the phenomena. On that account, remember, Proud and Detached are equally responsible for their country’s actions – their being so responsible is not up to them in any relevant sense. Why is it, then, that we want to distinguish between the flaws in their behavior when they claim not to be responsible for their country’s wrongdoing? How is it that Proud’s having taken pride in his country’s achievements makes a difference here, if it is not within Proud’s powers to change

¹⁰ Because Wolf does not consider the active nature of taking responsibility – and perhaps also because she thinks of things here primarily in aretaic terms (in terms of the nameless virtue of being willing to take responsibility, roughly), she cannot deliver this result, she cannot distinguish as clearly as the suggestion in the text between Proud and Detached.

the relevant responsibility-facts? I am not saying that nothing can be said here. Perhaps there are other explanations of the asymmetry between Proud and Detached – perhaps, for instance, on top of his other moral problems (also present in Detached), Proud is also hypocritical. But the most natural, unified explanation seems to be the former one, in terms of Proud’s having exercised – by taking pride in his country’s achievement – his normative power to make himself responsible for its wrongdoings.

We are now also in a position to solve a puzzle mentioned earlier. Prior to and independently of taking responsibility for your teen-aged son’s action, we’re here assuming, you are not responsible for his action. And yet, earlier I suggested that in some circumstances, if all you do to respond to certain relevant reactions is to say something like “But it is not my action; I am not responsible for it and its consequences.” you behave in a morally suspicious way. Why is this so, given that what you say is strictly speaking true?¹¹ The answer, it seems to me, is that sometimes by uttering this truth, you will be conveying your unwillingness to take responsibility, as you should. Hence the wrongness. Think again about the example of the sick friend worrying about the future of his daughter. If when he expresses his anxiety you respond merely by offering the (true) prediction that someone will take care of the girl, you behave wrongly, because by (merely) saying this you convey your unwillingness to promise to take care of the child, as you should. Hence the wrongness. The two cases are, I think, parallel.

5. The Responsibility Taken

But it is now time to address a worry that may have been there in the background for a while. What is, exactly, this responsibility that you are taking? What is the nature of this moral status, such that it can

¹¹ A possible answer – different from the one I am about to give in the text – is that you are here under a moral duty to behave (and feel) *as if* you are responsible (even though you know that you are not). I do not have a general objection to such duties to pretend. But they do call for explanation – why is it, in other words, that you have the duty to pretend here? Without such an explanation, the duty to pretend seems objectionably mysterious, and I don’t know how to explain it. Furthermore, I think the explanation that follows in the text is superior to one in terms of a duty to pretend also on phenomenological grounds.

be changed by an act of will in the described way? And what other implications follow from your being, or your failing to be, responsible for something in the sense relevant here?

The term “responsibility”, as it is widely noted, is multiply ambiguous. Let me start, then, by ruling out the kinds of responsibility I have not been talking about.

We can distinguish between moral and legal responsibility (each of many different kinds). My discussion, of course, was all in terms of moral responsibility. I will have nothing to say about legal responsibility here. But this does not mean that nothing of interest can be said in this context about legal responsibility. Perhaps, for instance, sometimes assuming legal responsibility for something can be a way of taking moral responsibility for it, or for something else. But I will not pursue such possibilities here.

Sometimes we use “responsible” to attribute a property to an agent, or perhaps an ability, or some such. Thus, we can ask what distinguishes responsible agents from creatures who are not responsible. This may or may not be shorthand for asking what distinguishes those who are responsible *for* their actions from those who are not. And we can say about someone that she is a very responsible teacher, and perhaps also that she is a responsible person, without specifying anything *for* which she is responsible. This too is a use of the term “responsibility” (and related terms) which is not our topic here.

Sometimes we use the term “responsibility” as standing for a purely causal relation. “The weather was responsible for the quality of the game” may be understood as the claim that the weather was a cause of the quality, or perhaps even *the* cause, or perhaps the one among many causes that we should direct our attention to¹². And because people and their actions may play causal roles as well, we can apply the term “responsibility” in this sense to them as well. But this too, of course, is not what we are interested in. (And clearly, one cannot by an act of will make it the case that one is responsible in this sense to an event in the past – this really *would* be backward-causation.)

¹² For this interesting suggestion see Baier (1972, 118).

I have already mentioned Baier's task-responsibility, or – perhaps somewhat more generally – the sense of “responsibility” in which to say that you are responsible for something is, roughly, to say that you are under a duty to bring it about, or to look after it, etc. This is a forward-looking sense of responsibility, and we are here interested in backward-looking responsibility, the kind of responsibility your son may bear to his (past) crime, and that arguably you may bear to it as well, once you take responsibility for it. Notice that it is some forward-looking sense of responsibility in which you can take responsibility for things that lie entirely outside the scope of your agency. Perhaps, for instance, a (morally and legally binding) contract for a camping event can allocate risks regarding the weather, so that the camping guide will have to return your money if it rains. One way of describing this case would be that the tour guide accepts or even takes responsibility for the weather. But this, to repeat, is not the sense of “responsibility” (and of “taking responsibility”) in which I am interested.

A natural alternative to all of these notions of responsibility – and a *very* natural attempt at capturing backward-looking responsibility – is an understanding of responsibility that ties it very closely to blameworthiness and praiseworthiness. Thus, you may think of responsibility in this sense as whatever it is that has to be added to wrongness to get to blameworthiness, or you may think of responsibility in this sense as that which is common to both praiseworthiness and blameworthiness. In this sense, an agent A is responsible for X if and only if, roughly speaking, if X is morally wrong or bad then A is blameworthy for X, and if X is morally right or good then A is praiseworthy for X. And it is, after all, in this sense (perhaps among others) that you are responsible for things in the core scope of your agency.

It is very important for me to emphasize, then, that this too is a sense of “responsibility” I am not talking about when putting forward my suggested understanding of taking responsibility¹³. To see

¹³ Adams (1985) ties his talk of responsibility and of taking responsibility (for “involuntary sins”) to blame, so it may seem as if here too there's a big difference between us. But in fact I do not think that this is so, because Adams is working with a *very* weak notion of blame: see Adams (1985, 21).

this point, it will be helpful to use a metaphor that is common in the related literature – that of a moral taint, or what goes on one’s moral ledger¹⁴. On this common metaphor, being blameworthy for something consists of being assigned a debit – on account of that thing – on one’s metaphorical moral ledger, and being praiseworthy for something amounts to being assigned a credit on that ledger on account of that thing. Of course, this is merely a metaphor, and it is not clear what exactly the metaphor is supposed to capture. But for my purposes here we do not need to worry about the literal truth in the vicinity here. The important point in our context can be put metaphorically: No later act of will can change the earlier entry on one’s ledger. If your country’s action does not taint your moral record (because it is not your action, or one for which you are responsible, as we’ve been assuming), if it does not constitute a debit on your moral ledger, then there doesn’t seem to be anything you can do at a later time to change that. Of course, there is plenty you can do at a later time to introduce *further* debits on your ledger. But there is nothing you can do to introduce such a debit for something for which you already failed to receive this debit. It is primarily this kind of responsibility to which the worry mentioned in the previous section does indeed apply – it is in this sense of “responsibility”, the one that is closely associated with praise and blame, that whether or not you’re responsible for something is just not up to you. So my suggested understanding of taking responsibility – according to which your taking responsibility for something makes it the case that you are responsible for it – cannot plausibly apply to this kind of responsibility. You can only be responsible in the blame- and praise-related way, then, to things that lie in the core of your agency¹⁵.

At this point you may be worried that no sense of “responsibility” remains for my account of taking responsibility to apply to. Let me shift to a more positive mode, then, and present the kind of

¹⁴ See, for instance, Zimmerman (2002, 555).

¹⁵ At times I talk about taking responsibility in terms of incorporating it into one’s agency. I do not intend this to mean that you thereby incorporate whatever it is you’re taking responsibility for into your *core* agency. Even after taking responsibility for something, you’re still not blameworthy for it (unless you are independently of the responsibility-taking).

thing I have in mind by “responsibility” here. I will use as my starting point Marina Oshana’s (1997) understanding of responsibility as accountability¹⁶. On this account,

... the judgment “X is responsible for Y” will express the beliefs “X did Y (or is Y), X possesses the qualities of a responsible agent, and X is accountable or answerable for Y.

And

“X is accountable for Y” can be unpacked as “It is appropriate that X explain her intentions in doing (or being) Y”. (77)

The main thought here is rather simple: Responsibility is sometimes not about what is attributable to whom, exactly, but about what kinds of explanations are appropriate, and in particular, whether it is appropriate to expect someone to explain an action (or some such thing).

Let me immediately mention one respect in which my ambitions are very different – and are much more restricted – than Oshana’s. Oshana is after a general account of moral responsibility, including that relevant for blame and praise. And she argues for the accountability understanding of responsibility as a candidate for such a general account (she is particularly interested in arguing that it is preferable to a Strawson-like reactive-attitudes account). I have my doubts about the success of this more general project¹⁷, which is anyway not mine. What is important – and sufficient – for my purposes is that there is *a* sense of responsibility that Oshana’s account (or something close to it) succeeds in capturing. And this claim seems plausible, I think, even if Oshana’s more general project cannot ultimately succeed.

¹⁶ Oshana develops here themes she finds in Baier, and possibly elsewhere too.

¹⁷ Here’s the main one: Oshana (following Baier) thinks that if you are responsible in the sense of having this obligation to explain, and you fail to explain, you are then blameworthy. I do not think this is a plausible thing to say. But I digress.

I also think that we can improve somewhat on Oshana's way of putting things¹⁸. First, Oshana speaks in terms of an explanation being appropriate. And at times (76) it sounds as if the question – whether an explanation is appropriately expected – is almost epistemological, depending on whether someone is likely to have an explanation available to her, or some such. But this is not, it seems to me, the right understanding of the appropriateness of explanation here (as I think Oshana herself understands elsewhere (77)). Furthermore, Oshana's account neglects the directionality or the second-personal nature¹⁹ of the phenomenon here. Her account of accountability does not include – not explicitly, anyway – anything about the explanation being not just appropriate but *owed*, and furthermore owed *to* someone in particular. But it seems to me that the kind of responsibility in the vicinity of what Oshana puts her finger on has precisely these features.

Because of these facts – the relevant appropriateness is not epistemic; the thing which is appropriate may be owed; and it may be owed to someone – it seems to me more fitting here to speak of an *apology* rather than an explanation. The sense of responsibility I am interested in, and the sense in which you can sometimes make it the case that you are responsible for something by taking responsibility for it, is the sense that makes an apology appropriate. Thus, somewhat roughly: A is responsible (in this sense) for X to P if and only if, if X is wrong or otherwise morally problematic, A owes an apology to P on account of X. This does not mean that Oshana's explanations are not relevant here. Apologies, I think, are the paradigmatic example of a wider phenomenon, one that may very well include Oshana's explanations (again, understood non-epistemically). Other related responses may include offering a justification, or an excuse. And thus we get, still somewhat roughly: *A is responsible (in this sense) for X to P if and only if, if X is wrong or otherwise morally problematic, A owes it to P to justify*

¹⁸ The quote above includes two parts that are irrelevant for my purposes: Because I am assuming that in all the relevant cases the relevant agent satisfies the conditions for general moral responsibility (as an ability), we do not need to worry about that. Also, because I do not attempt to answer here the question of what categories of things we can be responsible for, I don't need to worry about the "doing or being" in Oshana's characterization.

¹⁹ I borrow this way of talking here – though not anything of more substance – from Darwall (2006).

X (or the relevant *X*-related thing), or to offer an excuse for it, or to apologize for it, or to explain it, or something of this sort²⁰. And if all is well, this understanding of responsibility should sound familiar to you – it corresponds, I think, to a sense we employ all the time.

Let me emphasize again that it is in no way a part of what I am arguing for that this is *the* right understanding of responsibility. This is also why I settle for an if-and-only-if claim above: It is consistent with my use of this understanding of responsibility that the biconditional is not explanatorily basic, that there is a more basic truth here about the nature of responsibility that explains the biconditional. Indeed, it is even possible (and consistent with my use of the biconditional) that in these cases one owes an apology (say) *because* one is responsible in some more basic way²¹. All I will be relying on is that there is a sense of responsibility for which this biconditional holds. And it is in this sense of responsibility that you can – and sometimes should – take responsibility.

Let's return now to the three cases that have been with us since the start, and see how this understanding of responsibility naturally applies to them. You do not, it seems to me, owe an apology for your country's actions (to the victims, say), simply in virtue of your citizenship. Something else is needed. But you may be required to take responsibility. And if you do take responsibility, it will then be, under certain conditions, appropriate to expect you to offer a justification for your country's actions, or an excuse, or in the absence of either perhaps an apology. You owe such a response because you are now responsible (in this sense) for your country's relevant actions. And you are now responsible for them because you've taken responsibility for them (as you may have been morally required to do). Similarly, it seems to me, for the other two cases. Even if your driving was impeccable, you should take responsibility for the consequences of your action. This does not mean that you should think of yourself as blameworthy for them – you are not, and so you should not (you are not morally required to have

²⁰ I thank Adam Hochroth for correcting some errors in my characterization of this kind of responsibility.

²¹ I am not sure I understand Strawson's (1962) influential discussion of responsibility and the reactive attitudes, but I am quite confident that to the extent that I do understand it, I do not see how it is supposed to be helpful. So let me note that the point in the text suffices, I think, to show how nothing here is at all Strawsonian.

this false belief, after all)²². But if you take responsibility, you then become responsible, and then you may owe “your” injured pedestrian some rudimentary form of explanation. You may, for instance, owe it to him to present evidence for your good driving, and also to offer your sympathy for his loss. And on my account, you do not owe him these things independently of taking responsibility for his injury. If you do not take responsibility, you are not responsible. But you are in violation of your duty to take responsibility. And similarly, *mutatis mutandis*, for the parent case as well.

Notice that even though there is something forward-looking about this kind of responsibility – it is about an apology (or some such) that is owed from this point on – it is not *just* forward-looking, and is not just an instance of task-responsibility or some such. Two considerations make this clear: First, the relevant forward-looking thing here has something backward-looking built into it. If you have a duty to explain or justify or excuse or apologize, this is a duty that is necessarily related to something in the past. In this respect, such duties are importantly different from a responsibility (that is, simply a duty), for instance, to pay compensation or, say, the hospital bills. Second, not just *any* duty to apologize (for instance) will do here. If I reliably promise you to end all misery if you apologize to the world about the length of your nose, you are morally required, I take it, to apologize about the length of your nose. But this doesn’t show that you’re responsible in the sense we’re after for the length of your nose. What is needed, roughly, is that an apology will be *appropriate*, or that it will be owed *for the right kind of reasons*²³. Unfortunately, I do not have up my sleeve a solution to the wrong-kind-of-reasons problem, here or more generally. Fortunately, I don’t need one. For my purposes it is enough that we know –

²² Wolf (2001, e.g. 117) repeatedly thinks of a variant of this example in terms of the degree to which it is appropriate for the driver to blame herself. But this creates awkward complications, as when Wolf has to divorce blaming oneself from believing oneself blameworthy (2001, 125) in order to accommodate the point in the text here. For this reason, then, I do not think that thinking about the example – and about agent-regret more generally – in terms of blaming oneself is helpful.

²³ Notice that in the example in the text the directionality element is also likely to be missing.

often, at least – to distinguish between right and wrong kind of reasons²⁴. And the fact that the distinction is relevant here shows that the kind of responsibility I am talking about is not just that of a duty, or task- or forward-responsibility.

I need to discuss a complication here²⁵. I suggested that the right thing to say about the country case, for instance, is that you're not responsible (in this sense) for its actions, but that because of the nature of the relation between you and your country, you're morally required to take responsibility, and then to apologize or some such. And I have also noted in the previous section that apologizing may amount both to an implicit way of taking responsibility, and to living up to the responsibility thus assumed. But an alternative suggestion is that the nature of the relation between you and your country *directly* makes an apology (or some such) appropriate. In this accountability-sense, the argument goes, you *are* responsible for your country's actions, regardless of whether or not you take responsibility. Now, I have been assuming that people are not responsible for their country's actions merely in virtue of their citizenship. But it may now be argued that what made this claim plausible was a *different* sense of "responsibility". Perhaps mere citizenship is not enough for blame-related responsibility. But why think that it is not sufficient for accountability or apology-related responsibility either? And if so, the act of taking responsibility becomes entirely redundant.

I want to acknowledge that this alternative account is not without plausibility. And the comparative appeal of the two accounts may differ from one case to another. (For instance, this alternative account seems more plausible when applied to the driver case than when applied to the country case, I think; I return to this point in section 7.) But two kinds of consideration count against the

²⁴ A deeper reason why I don't need for my purposes here a (local) solution to the wrong-kind-of-reasons objection: Perhaps I would have needed one, had I put forward the biconditional unpacking the notion of responsibility relevant here as explanatorily basic. As explained above, though, I do not, and so everything I say here is consistent with the biconditional holding in virtue of some other thing, perhaps a more basic account of responsibility. And this more basic account may not be even initially vulnerable to the wrong-kind-of-reasons objection.

²⁵ I thank Peter Cane for relevant objections.

alternative account as a global one: First, why would anyone be expected to apologize (for instance) for actions that are not his or hers, and that are not under his or her control? To me, this seems already problematic. I acknowledge, of course, that this is hardly a conclusive or uncontroversial reason to reject the alternative account. But the guiding intuition underlying my suggested account still applies: According to my suggestion, you cannot be responsible (in this sense) for anything outside the core scope of your agency unless by an act of will you make yourself so responsible. This seems to me to be a feature of a theory of responsibility worth maintaining. Second, as I suggested above, there seems to be something active about the responsibility relevant in these cases, and this is a point nicely captured by my suggested account, but not by the competing account. This was a part of the point of introducing the comparison of Proud and Detached. If you agree that there's a difference between them regarding the extent to which they are responsible for their country's wrongdoings, then you already agree that the alternative account – one that considers the acts of will of the relevant agents redundant here – is not the full story of responsibility in such cases. And this, I think, is all that I need. And remember, finally, that I am here more in the business of establishing a possibility than an actuality – even if at times the alternative account works well enough, still my suggested account occupies a legitimate and possibly interesting position in logical space.²⁶

²⁶ Let me flag three further issues here, issues that I do not know what to say about: First, I have been talking about responsibility only for bad things. But we are often responsible for good things as well. It's not clear to me how – or even whether – what I say here can be applied to the positive case as well. Perhaps this is so because the notion of taking responsibility is only useful in negative contexts (I think). Second, I have been talking both of the power to take responsibility (so that you can take responsibility for some of your country's actions, but not for the motions of the planets), and about the duty or moral requirement to take responsibility. But the only cases I described were cases in which the power and the duty were both present. But it seems reasonable to assume that there are cases that lie within your penumbral agency so that taking responsibility for them is within your power, but that you are not morally required to take responsibility for them. There may be, in other words, cases of morally optional responsibility-taking. I do not know what to say of those. Third, it is interesting to ask about the revisability of taking responsibility. If I am now responsible because I have taken responsibility, can I do anything to undo that taking? I don't think there is a general answer to this question, not more general than the answers we can give to the analogous question about promises (once obligated because having promised, can one undo the normative effect of the promise?). But even if there is such a general answer, I cannot look for it here.

6. Penumbra Agency

I started my discussion with three examples of cases that lie – so I suggested – in the penumbra of an agent’s agency, things that the agent is not responsible for, but that she can become responsible for by taking responsibility for them. I have also insisted that in some such cases one plausibly has a moral duty to take responsibility, and thereby become responsible. And I have been insisting that I am here arguing merely for a possibility result, and that not much depends on whether you are convinced by these examples. But can anything more general be said here about the duty to take responsibility, and the related phenomenon of penumbral agency? When do we and when do we not have this moral duty, and why?

My discussion here is going to be (even) more incomplete than in the rest of this paper. But the intuitive thought seems to be that there are ways in which we are morally required to think of ourselves. In some – perhaps most – circumstances a parent should think of himself as a parent, he should take this true description of himself as important, normatively relevant, reason-giving. One of the things thinking of oneself as a parent (in this normatively rich way) includes is the willingness, even the determination, to incorporate things like one’s children’s actions into one’s agency. If you don’t do that (in the appropriate circumstances), then you fail to live up to the duty to think of yourself as a parent. Of course, we cannot force you to think of yourself as a parent. And if you do not, events like your children’s actions that lie in the penumbra of your agency will not be ones for which you are responsible (in the sense explained in the previous section). So the mere fact that these *are* actions of your children does not suffice for your being responsible. But if you do not incorporate them into your agency by taking responsibility for them, you are not thinking of yourself as a parent (in the normatively rich way needed here). And – being a parent – you *should* think of yourself as a parent. Hence the moral failure.

Similarly, you (arguably, and in certain but certainly not all circumstances) ought to think of yourself as your country’s citizen. And one of the things this means is that you should incorporate its

actions – lying in your penumbral agency – into your agency by taking responsibility for them. But if you do not think of yourself as your country’s citizen, you are not responsible for its wrongdoings. It’s just that if the circumstances are such that you *should* think of yourself as a citizen, the moral failure is precisely in failing to do so, in failing to take responsibility for your country’s actions.

The case of responsibility for the consequences of your actions is less clear, I think, but is still plausibly analogous to the two other cases. For arguably, we should take interest in the world²⁷, not just in our will or intention, and so we should think of ourselves as creatures whose intentions and bodily movements have effects in the world. So we should incorporate the consequences of our actions (some of them, that is, in some circumstances) – events lying in our penumbral agency – into our agency by taking responsibility for them. Furthermore, perhaps we are also morally required to think of ourselves as *limited*, as creatures who have some but far from full control over the consequences of their actions in the world²⁸. If so, we have this further reason to incorporate the unintended and unforeseen consequences of our actions into our agency. Indeed, it may be argued that we – all of us, pretty much – already did. For much like Proud – and much unlike Detached – we *do* take pride in and claim credit for the consequences of our actions and projects when they turn out well, as we intended them, and sometimes even when they turn out *surprisingly* well, in ways we did not intend or foresee that they would. If in the political case it is plausible to think of Proud’s taking pride in his country’s achievements as implicitly incorporating its actions into his agency and taking responsibility for them, a similar point can be made about the relation between each and every one of us and the consequences of his or her actions.

In general, then, I suggest that the related phenomena of penumbral agency and the moral requirement to take responsibility for things one would not be responsible for had responsibility not been taken are both the upshots of there being ways in which we are morally required to think of

²⁷ This is a central point in Wolf (2001). For a related point, see Adams (1985, 16).

²⁸ A conversation with Bob Kane helped me to see this.

ourselves in a normatively rich kind of way. If you like identity-talk, you may want to put this point by saying that there are moral requirements that apply to the way in which we construct our identities, and indeed there are some things we are morally required to construct into our (normatively significant) identities. If you like self-talk, you may want to put the point by saying that there sometimes are moral requirements to incorporate certain things into one's self. But regardless of your favorite jargon, the point remains the same: We have the (limited) power to incorporate certain things into our agency by (perhaps among other things) taking responsibility for them, and sometimes – when doing so is called-for by a way we ought to think of ourselves in this normatively rich kind of way – this is precisely what we should do.

There is, of course, something impressionistic about this section, and it is not as precise as I would want it to be. But it does, I think, capture what we naturally want to say about cases like the three I have been using as my examples throughout. It may even throw some light on the questions in these contexts regarding which we do not know what exactly we want to say. Perhaps, for instance, the (or at least an) important question about moral luck and responsibility for consequences is not in terms of blameworthiness, and not about whether we *are* responsible, but about whether we ought to take responsibility for the consequences of our actions (and perhaps we already have done so). And given the discussion in this section, perhaps this means that the question we should be interested in here is precisely whether we should think of ourselves – in the normatively rich way in which those of us who are parents should presumably think of ourselves as parents – as creatures who bring about only partly controlled changes in the world²⁹.

Furthermore, once we notice the room in logical space for the phenomenon of taking responsibility as characterized in this paper, and once we notice its relation to the phenomenon of

²⁹ There are hints at this way of going about the moral luck debate in Joseph Raz's contribution to this volume. But there are also things in Raz's treatment of agent-regret with which I rather strongly disagree. I cannot pursue them here.

penumbral agency and its normative significance, we see further possible examples almost everywhere (this, at any rate, is what has happened to me since I started thinking about these phenomena). Thus, perhaps the way to make progress on a theory of distributive justice that is luck-egalitarian at least in spirit is to be sensitive not just to the distinction between disadvantages people are responsible for and those for which they aren't (those that are, as it were, a matter of luck), but also to the further category of disadvantages for which people are not responsible, but for which they are morally required to take responsibility. And perhaps the thing to say about the relation between me today and actions in my distant past is not that I am – as I am today – responsible for them, but that I should still take responsibility for them. At least, the discussion of taking responsibility serves to show that there are more possible ways to go in these and other cases than we may have previously thought. And this too, I think, is not without value.

7. Returning to Moral Luck

This concludes, then, my general discussion of taking responsibility and penumbral agency. But I now want to return to the topic I started with, and note, in conclusion, what follows from the discussion of taking responsibility to the issue of moral luck (more broadly than as exemplified by the driver example).

For reasons largely independent of anything in this paper, then, I stand by my rejection of any kind of moral luck. I believe, in other words, that necessarily, if two agents are alike in what is under their control they are also alike in the extent to which they are praise- or blameworthy, and indeed morally responsible (for the relevant thing). But I think that such agents may differ in what they are morally required to take responsibility for. And then, of course, if they do, they will differ in what they are responsible for (in the sense developed in section 5), in virtue of having differentially taken responsibility. And I indicated the way in which the discussion should proceed about who should take responsibility for what – though I remain largely agnostic on what the consequences of this discussion

are likely to be, and in particular whether we are – and if so, when we are – required to take responsibility for the consequences of our actions (and perhaps also for the circumstances in which we happen to find ourselves, or for our character traits, if we want a to pursue a similar line on the phenomena Nagel (1976) calls circumstantial and constitutive moral luck).

Getting back to Williams on agent-regret, then, we get a vindication, I think, of Williams' insight, but not of the conclusions he was (apparently) trying to draw from it³⁰. True, agent-regret is deeply important. And there indeed *is* something problematic – morally and otherwise – about someone who would fail to feel agent-regret. But all of this is not indicative of blame, or blame-related responsibility, or even accountability-type responsibility that is already there. Rather, all of this is indicative of a moral requirement to take responsibility that the agent violates. Feeling agent-regret, I want to suggest, is (sometimes) an appropriate response to the duty to take responsibility. It may be appropriate in that it is a way of acknowledging the existence of this duty. Or perhaps it is the appropriate response not so much to the duty to take responsibility, but to one's being responsible in virtue of having taken responsibility – for as explained above, when it comes to the consequences of our actions, we are all analogous to Proud rather than to Detached from the political example.

You may be worried, though, that when all is said and done, this way of accommodating the agent-regret intuition apparently without conceding the existence of moral luck amounts in effect to precisely this concession. For on this suggestion, while there is no difference in blameworthiness (etc.) between two equally negligent drivers when one of them hit a pedestrian and the other didn't, still there is a difference between them regarding what they are morally required to take responsibility for. And this difference too is here a matter of luck – because a pedestrian happened to jump in front of the

³⁰ Wolf (2001, 126) presents her closely related observations as leading to a kind of a compromise position in the moral luck debate. But – as I think is implied by Wolf's own tone in the paper in general – the description in the text here is more precise. Perhaps it is harder for Wolf to present things more explicitly in this way because she does not note the remaining difficulty I am about to raise, and she doesn't use the distinction between moral luck and morally significant plain luck – a central distinction, in my opinion, on the way to viewing the main points in this paper as taking sides on the moral luck debate.

car of one of these drivers but not the other, the former is now morally required to take responsibility for something the latter is not morally required to take responsibility for. Isn't this just as bad as accepting moral luck? Indeed, isn't this just another case of moral luck?

It is a case of luck, alright, but not moral luck in the sense we've been talking about. For on this account, what is sensitive to luck is not (primarily) what you are blameworthy for, or what you are (backwards-) responsible for. Rather, what is sensitive to luck is which moral duties apply to you – in particular, whether you are morally required to take responsibility for something. And perhaps this phenomenon – the luck-sensitivity of one's moral duties – is itself interesting, perhaps even theoretically problematic. But it cannot be denied that it exists. Clearly, whether you are morally required to give to famine relief depends on whether you find yourself in circumstances in which people suffer from hunger, whether you are morally required not to move your fisted hand rapidly forwards depends on whether you find yourself in circumstances in which someone stands right in front of you, and so on. It cannot be seriously doubted, that is, that morality is in these trivial ways sensitive to circumstances. What you are morally required to do depends on the circumstances you are in, and many of those are not under your control, they are for you a matter of luck. So it is important to distinguish between (purported) cases of moral luck on one side, and cases of plain luck that has moral implications, or morally significant plain luck on the other. And while the latter may also be problematic, it is not as problematic as the former, and anyway its uncontroversial existence does not entail the highly controversial existence of the former.³¹

Thus, acknowledging the role of a moral duty to take responsibility allows us to accommodate Williams' insight by classifying the case of agent-regret as a case of morally significant plain luck. And so we do not need, in order to accommodate it, to accept the much more radical and problematic thesis of

³¹ I discuss in some detail the distinction between moral luck and morally significant plain luck in Enoch (2008).

moral luck. That his own insights can lead us in directions so different from those in which they led him is yet another proof of Bernard Williams's philosophical stature.

References

- Robert Merrihew Adams, "Involuntary Sins", *Philosophical Review* 94 (1) (1985), 3-31.
- Kurt Baier (1972), "Types of Responsibility", in Peter A. French Ed., *The Spectrum of Responsibility* (New York: St. Martin's Press, 1972), 117-122.
- Kurt Baier (1987), "Moral and Legal Responsibility", in Mark Siegler (ed.) *Medical Innovation and Bad Outcomes: Legal, Social, and Ethical Responses* (Ann Arbor, Health Administration Press, 1987), 101-129.
- Stephen Darwall, *The Second-Person Standpoint: Morality, Respect, and Accountability* (Cambridge: Harvard University Press, 2006).
- Enoch, David. 'Luck between Morality, Law, and Justice.' *Theoretical Inquiries in Law* 9 (2008), 23-59.
- David Enoch (manuscript), "Giving Practical Reasons".
- Enoch, David and Marmor, Andrei. 'The Case against Moral Luck.', *Law and Philosophy* 26 (2007): 405-436.
- John Martin Fischer and Mark Ravizza, "Précis of Responsibility and Control: A Theory of Moral Responsibility", *Philosophy and Phenomenological Research* 61(2) (2000), 441-5.
- Marina A. L. Oshana, "Ascriptions of Responsibility", *American Philosophical Quarterly* 34 (1997), 71-83.
- Marion Smiley (2005), "Collective Responsibility", *Stanford Encyclopedia of Philosophy*, available here: <http://plato.stanford.edu/entries/collective-responsibility/>
- P. F. Strawson (1962), "Freedom and Resentment", *Proceedings of the British Academy* 48, 1-25.

- Nagel, Thomas. 'Moral Luck.' *Proceedings of the Aristotelian Society*, suppl. 50 (1976): 137–152;
reprinted (in a revised version) in his *Mortal Questions*. Cambridge: Cambridge University Press,
1979, pp. 24–38; and in Statman (1993), 57-71.
- Williams, Bernard. 'Moral Luck.' *Proceedings of the Aristotelian Society*, suppl. 50 (1976): 115–136;
reprinted (in a revised version) in his *Moral Luck*. Cambridge: Cambridge University Press, 1981.
20–39; and in Statman, 35-55.
- Bernard Williams, "Postscript", in Statman, Daniel. Ed. *Moral Luck*. Albany: State University of New York
Press, 1993, 251-8.
- Susan Wolf, "The Moral of Moral Luck", *Philosophical Exchange* 31 (2000), 4-19; reprinted in Cheshire
Calhoun (ed.) "Setting the Moral Compass: Essays by Women Philosophers" (Oxford: Oxford
University Press, 2003), 113-127.
- Zimmerman, M. (2002), "Taking Luck Seriously", *The Journal of Philosophy* 99, 553-576.