

## Agency and Luck

Joseph Raz<sup>1</sup>

Williams's 'Moral Luck' proceeds indirectly. He points out that widely held beliefs about what we are responsible for, and about differences in degree of blameworthiness, are inconsistent with a certain conception both of responsibility and of moral blame. If we hold on to our beliefs about when we are responsible, while understanding blame and responsibility according to that conception, we are committed to the existence of what is, according to that conception, moral luck, whose existence is antithetical to that conception of responsibility and blame. He did so to persuade us to abandon those conceptions of responsibility and of blameworthiness. But his paper did not advance an alternative. Rather, his paper aims to contribute to the vindication of common beliefs about acts and consequences for which we are responsible and about degrees of blame, by explaining an ineradicable feature of our attitude to the consequences of our actions. Critics have rightly pointed out that he has not established the relevance of that feature to questions of responsibility. The connection is indirect. The article implies that that attitude to the consequences of our actions is presupposed by a sound conception of responsibility. But Williams did not tell us what that conception is.<sup>2</sup>

I will proceed in the opposite direction. After singling out the concept of responsibility my paper is about, and offering a first step towards its explanation (Section 1), I will turn to the ways we are attached to the consequences of our actions, and to Williams's views about them. My views are not his. But the paper is not only indebted to his inspiration. It is true to his way of exploring the phenomena discussed. Or so I hope.

---

<sup>1</sup> Professor, Columbia University, NYC. I am grateful for discussions and comments from David Enoch, Penelope Bulloch, Andrei Marmor, Ulrike Heuer, Rebecca Prebble, and Nandi Theunissen.

<sup>2</sup> He did at least partly explain it in later work, especially in SHAME & NECESSITY. But I will not follow those views of his.

## I. RESPONSIBILITY

Given the diversity of inter-related uses, or senses, of 'responsibility', a few essentially terminological observations are in order: Naturally we are not interested in the use of 'responsibility' to indicate causality (as in "the earthquake is responsible for the power failure"). Similarly, neither the use of 'is responsible' as a commendation (as in "you can trust him. He is a responsible fellow"), nor its reverse, namely its use as equivalent to blameworthy (as "the doctor is responsible for his death" would normally be understood) are of interest. We will have to deal with the conditions under which people are blameworthy, but – to avoid confusion – we will not use 'responsible' to designate blameworthy. Nearer to our concern is the sense of 'responsibility' which is close to that of a duty (as in "it was your responsibility to secure the building at the end of the evening").

Clearly not any sort of creature can have responsibilities in that sense, only responsible creatures can. Here being responsible indicates possession of some capacities. People (and other beings) are responsible<sub>1</sub> if and only if they have the capacity for rational agency (e.g. "he is not in his right mind and therefore not responsible for his actions").<sup>3</sup> Responsibility<sub>1</sub>, or capacity responsibility as – following Hart – I will call it, is the core notion, not etymologically, but in being presupposed by the others. Our focus is on responsibility for conduct, and people are responsible<sub>2</sub> for performing an action only if they are responsible<sub>1</sub>, and their action is appropriately related to their powers of rational agency. People are not responsible<sub>2</sub> for actions that they performed if, for example, their powers of rational agency were temporarily suspended or disabled by hypnosis or heavy sedation.

Three obvious necessary conditions of responsibility<sub>2</sub> for some X are: First, that one brought X about, preserved it, or contributed to its coming about or being preserved, or that one allowed it to continue.<sup>4</sup> (Most of the time I will inquire about responsibility<sub>2</sub> for actions, assuming that if one is responsible<sub>2</sub> for a state or an event,

---

<sup>3</sup> I will ignore the possibility, if possibility it is, that there could be creatures of species which are incapable of action, but capable of thought (and of mental actions), and the ways they may be responsible for their thoughts and mental actions.

<sup>4</sup> This condition raises the question of the degree of proximity and the character of the causal role required for its satisfaction, which will not be considered here.

etc. that is in virtue of responsibility<sub>2</sub> for an action which brought it about, caused it, preserved it, or allowed it to stay in existence, etc. This greatly simplifies exposition, and whatever distortion it introduces can be readily corrected.) Second, that one has powers of rational agency; and third, that they were not suspended or disabled at the time. The three are not, however, sufficient. The third condition indicates the need to preserve a connection between responsibility<sub>2</sub> and the powers of rational agency. But it does not itself establish a positive connection. It merely affirms that in the circumstances of the action a connection was possible.

The three conditions make clear that responsibility<sub>2</sub> is a matter of degree. Or rather, that responsibility<sub>2</sub> can be used either as an on/off concept, requiring that the conditions are met at some threshold level or better, or as admitting various degrees. To simplify the discussion I will treat it as an on/off concept.

What, then, is the relationship between capacities and actions which must obtain for agents to be responsible<sub>2</sub> for their actions? One initially appealing answer would have it that we are responsible<sub>2</sub> for  $\phi$ ing only if our  $\phi$ ing was guided and controlled by our powers of rational agency. What does this **Guidance Principle** amount to? An action is so guided and controlled if and only if, first, it is either done for (what the agent takes to be) a sufficient reason, or is done, knowing what one is doing (thus implicitly accepting that if there is a case against it, it does not defeat the case for it)<sup>5</sup>, by doing another action for (what the agent takes to be) a sufficient reason, and, second, in doing it one is not doing anything else which one believes that it would be better not to do.

The thinking behind this principle is readily explained. When we initiate an action because we see, as we believe, a reason for it, and whatever consequences we believe it will or probably will have are, if not desirable in themselves, at least not sufficient to make the action undesirable to us, and when in performing it we do not also perform other actions we would rather not perform, like stumbling and injuring other people, then our powers of rational agency are in charge, and we initiate and

---

<sup>5</sup> That agents take it that there is no sufficient case against the action is not so much an implication of the fact that they know what they are doing as a clarification of what level of knowledge is required: such as to imply that if they perform the action they believe that the case against it does not defeat the reasons for it.

guide the action by their use. In such cases the action manifests our powers of rational agency and we are responsible for it.

The Guidance Principle does not take people to be responsible<sub>2</sub> for all their intentional actions, for some intentional actions are neither done for a reason, nor by performing another action for a reason. Nor does it take people to be responsible<sub>2</sub> for all the actions they do for what they take to be a reason, for the performance of some of these actions is not controlled by their agents. The principle never holds one responsible<sub>2</sub> for accidental actions whose performance or consequences were not foreseen, nor for any omissions, other than those which were either decided upon or foreseen (intentional omissions).<sup>6</sup>

The appeal of the Guidance Principle is in relating what we are responsible<sub>2</sub> for to our powers of agency in the right way: we are responsible<sub>2</sub> for actions taken and guided by our powers of agency. By mixing two components: acting for what one takes to be a reason, and control, the Principle identifies actions regarding which we are maximally responsible<sub>2</sub>, in that where it applies we are responsible<sub>2</sub> for all aspects of the action of which we were aware. As a matter of fact we are also responsible<sub>2</sub> for some aspects of actions for which we are not fully responsible<sub>2</sub> accordance to the Guidance Principle. For example, as the **Intention Principle** asserts, we are responsible for the intended and for the foreseen aspects of actions for a believed reason even if we do not control some other aspects of the action of which we are aware.<sup>7</sup> We are also, given normal circumstances, responsible for weak-willed actions, even though the Guidance Principle does not establish that responsibility either.

It seems reasonable to take both the Guidance and the Intention Principles to state sufficient grounds for responsibility<sub>2</sub>. But their satisfaction is not necessary for responsibility<sub>2</sub>. We are responsible<sub>2</sub> for some unintentional omissions (e.g., in normal circumstances, for forgetting to do what we had to do) and for some accidental

---

<sup>6</sup> Another simplifying terminological convention I adopt here is to refer to the non-performance of any action on an occasion in which one could have performed it as an omission.

<sup>7</sup> As will be obvious, in the text I identify actions with events of a special kind, and aspects of actions as what is sometimes referred to as action under a description.

actions (e.g. for some accidental bumping into people while walking in the street). More generally, but not exclusively, we are responsible<sub>2</sub> for negligent acts and omissions. One way of denying negligence is to establish that one was not responsible<sub>2</sub> for the act or omission on which the charge of negligence is based. But neither the Intention nor the Guidance Principle applies to negligence.<sup>8</sup>

The problem is that the Guidance Principle makes responsibility depend on too tight a connection between our powers of agency and our actions. For us to be responsible for an action it must, according to the Principle, be successfully guided by our powers of agency. In fact we are responsible for actions where the guidance fails: when we accidentally bump into passers-by while walking, when we forget to do something we intended to do, when we act against our better judgement. The Intention Principle seems plausible, and extends to some cases of responsibility<sub>2</sub> which escape the Guidance Principle. But so far we saw no explanation for it, no account of how the fact that an action is intentional relates it to capacity responsibility to make us responsible<sub>2</sub> for it. Besides, even it does not cover all acts for which we are responsible, as the case of negligence illustrates.

The Guidance Principle makes responsibility turn on successful guidance. The success referred to is not that of doing what we have adequate reason to do, and avoiding what we have conclusive reasons to avoid – or something like that. It is not success consisting in following right reason. It is the successful functioning of our capacities of rational agency. People often do what they should not do without their powers of agency malfunctioning. This, for example, can be the case when they have false, but rational, beliefs about what they ought to do. We are used to distinguishing between failure in the functioning of our powers, be they powers to control our limbs, or mental powers called upon in forming and executing intentions, and failure in ‘getting it right’, in taking the right action. The way the distinction applies is controversial, and most likely partly depends on normative considerations. But the existence of the distinction is generally (often implicitly) recognised and underlies, or so I claim, responsibility<sub>2</sub>.

---

<sup>8</sup> It is possible to contend that the notion of negligence is incoherent, or that we can never be negligent. I will assume that neither is the case.

Conduct for which we are (non-derivatively) responsible<sub>2</sub> is conduct which is the result of the functioning, successful or failed, of our powers of rational agency. Let me dub this the **Rational Functioning Principle**. It applies to basic responsibility, and is to be supplemented by principles of derivative responsibility, e.g. an agent is responsible<sub>2</sub> for φing in conditions which would otherwise exempt one from responsibility<sub>2</sub> if they were created by that agent (in conduct for which he is responsible<sub>2</sub>) with the intention to exempt him from responsibility. There are other important principles of derivative responsibility<sub>2</sub>.

I have merely lightly sketched a conception of responsibility. It is to be judged by two tests. First, whether it really does apply to conduct for which we are responsible<sub>2</sub> and to no other. Second, whether the relations between powers of rational agency and action that it insists on are the right ones.

If the **Rational Functioning Principle** is sound we can expect the first test to be inconclusive. That is because it relies on several kinds of judgments whose soundness is independent of this conception of responsibility. First, there is the question whether one's powers of rational agency were temporarily disabled so that one is not responsible for any conduct performed during that period. Second, there is the question whether one's powers malfunctioned. For example, the fact that one does not rely in deliberation or action on all the propositions entailed by one's beliefs does not establish that one's rational powers malfunctioned. On the other hand, failure to realize and rely on some implications of one's beliefs does point to a malfunction of rational capacities. There is no determinate boundary between the two classes of case.

This makes it all the more important that there be a sound rationale behind this conception of responsibility<sub>2</sub>, and that is not at all obvious. We can understand why we are responsible<sub>2</sub> for conduct successfully guided by our rational powers, but why are we responsible<sub>2</sub> for conduct which is the result of a malfunctioning of those powers? This is where we turn to Williams for help.<sup>9</sup>

---

<sup>9</sup> I have discussed the question further in "Responsibility and the Negligence Standard", see <http://josephnrz.googlepages.com/agency%2Cresponsibilityandluck>.

## 2. AGENT-REGRET<sup>10</sup>

Williams, who gave the question of moral luck its name, did not focus on responsibility. Rather, he started by identifying one special attitude people may have to their actions and their consequences which we now know by the name he gave it: agent-regret. In doing so he signalled that crucial to understanding the question of moral luck is at least one way in which the consequences of our actions matter to us. I too will delay consideration of moral luck, preceding it with a reflection on the theme which occupied Williams.

*The constitutive thought of regret in general is something like "how much better if it had been otherwise", and the feeling can in principle apply to anything of which one can form some conception of how it might have been otherwise, together with consciousness of how things would then have been better. In this general sense of regret, what are regretted are states of affairs, and they can be regretted, in principle, by anyone who knows of them. But there is a particularly important species of regret, which I shall call "agent-regret", which a person can feel only towards his own past actions (or, at most, actions in which he regards himself as a participant). In this case, the supposed possible difference is that one might have acted otherwise. (123)*

This passage identifies agent-regret by its object: it is regret for having acted in some way. The regret is essentially self-referential. My regret that JR has done something is not agent-regret. My regret that I have so acted may be. Williams is clear that that is not the only difference:

*There can be cases directed towards one's own past action which are not cases of agent-regret, because the past action is regarded purely externally, as one might regard anyone else's action. (123)*

Agent-regret has, he explains, a specific expression, to do with a desire that one had not done what one did, and a desire to repair, to undo what one did or come as close to that as possible.

Arguably Williams's account of the feeling is incomplete, for it does not fully identify the difference between regretting one's own action which is agent-regret and

---

<sup>10</sup> This paper, and the second section in particular, is heavily influenced by B. Williams. However, my purpose is not exegetical. I aim to use his views and develop some aspects of them. I therefore took the liberty when describing Williams's views of paraphrasing in a way which is not altogether faithful to Williams's preferred formulations. Some of my formulations are cognitivist renderings of Williams's views. As this happens several times I did not alert the reader to the fact separately on each occasion. I consign to an Appendix observations on some of the aspects of Williams's views which I do not share.

regretting one's own action which is not.<sup>11</sup> I will suggest that an additional aspect of agent-regret accounts for that difference. But, as Williams does not mention this feature, I am not sure that the emotion that I am describing is the one he had in mind, though it seems to me consonant with his thought, and to mark an important distinction between two kinds of self-directed regret. The additional element, I suggest, is that agent-regret relates to one's sense of who one is. When I agent-regret an action of mine I feel bad or sorry about being or having become a person who acted in that way.<sup>12</sup>

According to my suggestion agent-regret is a feeling which can be associated with a number of distinct attitudes to oneself, united by being attitudes to who one is. Here is one imaginary example: I admire Dustin Hoffman, and would like to see him in the flesh. Being at home with a friend (who shares my attitude to Dustin Hoffman) I learn that he is nearby visiting someone in Montague Square, and that a few people are there waiting to see him come out. However, I do not go there and miss that chance to see him. Later I regret (a) missing the opportunity, and (b) not taking the very little trouble to go there to see him. Reflecting on my feelings I conclude that my regret about my failure is of the same kind as my regret that my friend did not go to see Hoffman. She and I are unlikely to have another opportunity to see him, and I regret that we missed the opportunity we had.

Given what Williams says about agent-regret, this is not agent-regret. It is a regret that I have had or did not have a certain experience, or that my biography

---

<sup>11</sup> See pages 125, 126. Williams points out that some desires to compensate do not involve agent-regret. That is the case when insuring against any harm done by the regretted action would have been taken to be sufficient to put an end to the regret. Nor does the desire to undo itself explain the essential self-referential character of agent-regret. I can regret your action, and desire to undo what you did, or to come as close to that as is possible. Similarly I can wish that some other people did not act as they did. The self-referential element is in the difference between wishing that JR did not act as he did and wishing that I did not act as I did. But what is the significance of that distinction (its meaning, not only its logical features)? I do not think that Williams explains that.

<sup>12</sup> Meir Dan-Cohen's discussion ('Luck & Identity' *Theoretical Inquiries in Law*, Vol. 9:1) in explaining Williams's *Gaugin* example as turning on Gaugin's decision being self-constituting, in that it relates to a pursuit central to his life from then on, lends some support both to my 'additional element', on which I rely in identifying the emotion, and to my contention below that Williams's explanation of the conditions under which the emotion is justified do not extend to simple cases of agent-regret.

does or does not include performing a certain action. It is not first-personal or self-referential in the right way. Why not? Suppose we vary the example by adding that I failed to go out to see Hoffman because I was ashamed of being seen to engage in star-gazing. In fact I find nothing wrong in moderate star-gazing, and my regret now includes feeling bad about myself for having succumbed to a feeling of shame which I regard as misguided and snobbish. If that is how I regret not going out then it is agent-regret. In this case I regret being a person who failed to go out.

But do I not also regret in the same way that my friend succumbed to an inhibition that she regards as unjustified (and let us assume that that is so)? I may well, but that regret is not self-referential in the way that my regret about myself is. I may hear a story – like the Hoffman story (and I mean just what was done on the occasion and the explanation, excluding the later regrets about it) – about someone, let's call him JR, and having forgotten that I behaved in that way myself I may or may not regret that JR succumbed to such an inhibition (there is nothing wrong in not caring enough about an unknown stranger not to regret that, indeed not to care whether, he is a person with this or that disposition). When I realise that I am JR I *cannot avoid* the regret, and it is a different kind of regret from the one I had before (if I had one before) about JR. It is regret that *I* am such a person. The essentially self-referential character of regret is particularly poignant due to its being, in part, about the person one is or was, as manifested on that occasion. It is poignant in being not regret that there is such a person, but that *I* am such a person. More specifically this instance of self-regret, though not all, involves something of a self-reproach, and self-reproach is essentially self-referential.<sup>13</sup>

The Hoffman example is of a case where the regret is motivated by the realisation, or the confirmation, of a known weakness. Other cases of agent-regret are different. In particular in some of them the agent regrets having become, through the action, someone he would rather not be. The person who runs over a child, through no fault

---

<sup>13</sup> Obviously, given the example I used, I do not propose that agent-regret can exist only regarding important actions, actions which embrace projects which give meaning to the life of the agent, as in the case of *Gauguin*. In such cases it may be particularly easy to see the connection between the person the agent is and his actions. But the connection is more general.

of his own, and kills him becomes a killer, someone who killed a child, and he regrets that. Having killed is something which may haunt him, and affect his attitudes to himself and to the rest of the world. (I mention these possible consequences of having killed a child to indicate the significance that being a person who killed has for some people. I do not mean that one agent-regrets the killing because one would rather not suffer these consequences.) Guilt and fault are not the only factors which can have such an effect on our sense of who we are. There is a genuine difference between agent-regret in which one's regret is directed at who one is, or who one has become through one's action, and other kinds of regrets about one's actions. And the difference is significant in that only in the one case does the agent regret being a person who so acts.

In drawing our attention to the prevalence of agent-regret Williams reminds us that we are attached to, care about, our involvement in the world, including aspects of it which are beyond our control. It will be evident by now that I do not think that agent-regret is unique in displaying the way our sense of 'who we are' is connected to our perception of what we do. Similar connections are displayed when we are proud, or ashamed, or feel guilty that **we** did this or that. Feelings of agent-regret serve as a good example which brings out the point, but other emotions will do as well.<sup>14</sup>

However, to vindicate this attachment Williams has to show that feelings of agent-regret can be justified, and to explain their significance in our lives. Williams provides a beginning of a reply to the first of these two questions:

*... it would be a kind of insanity never to experience sentiments of this kind ... and it would be an insane concept of rationality which insisted that a rational person never would. To insist on such a conception of rationality, moreover, would, apart from other kinds of absurdity, suggest a large falsehood: that we might, if we conducted ourselves clear-headedly enough, entirely detach ourselves from the unintentional aspects of our actions. (125)*

---

<sup>14</sup> Nor is there any reason to think that Williams would disagree here. He focusses on agent-regret because he finds in it a clear example of retroactive justification. It is less puzzling that our pride in an action depends on outcomes beyond our control than that the very justification of an action depends on it.

So long as we are the kind of creatures we are, he observes, we cannot detach ourselves from the unintentional aspects of our actions, a detachment necessary for the elimination of agent-regret. Therefore, it would be wrong ('insane') to think that such attachment is always irrational.

I will use 'irrational' of an action or an attitude to indicate that it is a product of a malfunction of our rational capacities. I assume that Williams is using it more broadly, to indicate that the attitude or feeling is against reason, that it cannot be supported by reason, meaning something close to saying that if it is not irrational then it is alright, not defective, and in that sense justified. He is saying that while sometimes a person's feeling of agent-regret is unjustified, there are no grounds for thinking that agent-regret cannot ever be justified, for that would leave us with the puzzle of why we cannot avoid a feeling which is necessarily unjustified, not even when we believe that it is unjustified.

This is not a conclusive argument. To supplement it we need to answer the second question I mentioned, we need to understand the significance of agent-regret in our life. If it is not only inescapable but also plays a significant role in our life then some instances of it can be justified. When is agent-regret justified? When its occurrence is appropriately related to the significance the feeling has for us. Here I part company with Williams, because his explanation applies only to regretting important, life-changing decisions. He expressed the hope that his discussion will illuminate more mundane occurrences, but I do not see how the rationale he offers for it can do that.<sup>15</sup> Here are examples of reasonable but minor cases of feeling agent-regret: I say something uncomplimentary to a friend hoping that it will spur him to confront his current difficulties. It may achieve its goal, and he may be grateful for my rudeness. But it may misfire, give offence and nothing more. In that case both of us may forget the incident before long. But before I do I may well regret my remarks, and that may be an appropriate case of agent-regret, at least it will be agent-regret by all the criteria which Williams mentions, as well as by mine. It will not, however, relate to a failure in a project which contributes to the meaning of my

---

<sup>15</sup> For his argument see the Appendix . It is far from clear that his explanation succeeds even in the cases to which he applies it.

life. My project, to prod my friend a bit, was purely altruistic, and its presence in my life was never meant to involve anything more than doing what I did. Williams fails to explain why agent-regret is here appropriate.

Other cases, while apparently cases of appropriate agent-regret, stray even further from the type of case he has in mind. Suppose that when I am on edge I express myself indelicately, in a way which could understandably and reasonably offend the friend whom I am addressing. I did not intend to offend, and do not feel that what I said accurately represents how I feel or what I believe. I regret having said what I said. I do not merely regret that JR said it; I regret that I said it. My regret is agent-regret. So far I said nothing about the consequences of what I said.<sup>16</sup> We can schematically distinguish three possibilities: (1) my friend does not take offence (perhaps his attention is on something else); (2) he is mildly offended, but quickly forgets the whole episode; (3) he is deeply offended and our relationship is diminished for a significant period, or declines altogether. To simplify let me assume that whatever his reaction it is not unreasonable or inapt. Only the third of these possibilities would warrant agent-regret on Williams's understanding of its justification. But it seems to me that I may well regret what I have done in all three cases (though possibly my regret may be greater in the third). Can such regret be justified? "My" kind of agent-regret, regret about who one is, is becoming or has become, comes with a ready explanation: we can reasonably wish or aspire to be like this and not like that, and so long as our regret at having become different is attached to such reasonable aspiration and is proportionate to its importance it is reasonable. Given how fundamental such emotions are to the kind of animals we are no other vindication is needed, and probably no other is possible.<sup>17</sup>

---

<sup>16</sup> Of course, my success in saying it is due, in part, to elements beyond my control. But they are not related (in the relevant way) to my regret.

<sup>17</sup> A clarification regarding justification: It is sometimes assumed that if an action or attitude is justified then not taking the action or not having the attitude is unjustified. But that is not generally true, and I will assume no such implication. Justification is permissive, as we might say. The justification of regret does not imply anything about the justification of its absence. Second, one can consider justification from a point of view, or a perspective, comprising a limited range of concerns. That is what we do when we talk of economic justification. Or one can take justification to be an all-things-considered judgment, and that is how it is taken here. The difficulty, not to be resolved here, is how to understand that notion. For example,

### 3. ENGAGING WITH THE WORLD

Williams discusses agent-regret to impress on us how deeply connected we are to some aspects of the world which are beyond our control. He did not, in that article, discuss responsibility. Is his discussion irrelevant to questions of moral luck as many have alleged? No, and yet - Yes. No, for agent-regret is neither a necessary nor a sufficient condition for responsibility. We are responsible for what we need not regret, and we properly regret actions and consequences for which we are not responsible. Yes – but very indirectly, for to understand responsibility we need to understand our attachment to consequences of our actions which are beyond our control, but the connectedness which governs responsibility is not that manifested by agent-regret.

To identify what that connectedness is we can start with the platitude that our sense of who we are is shaped through our life, which means through acting in the world and through our experiences of being in the world. But how does that bear on issues of responsibility? To answer, it is necessary to challenge the identity between matters beyond our exclusive control and matters of luck. Through their experience in life people develop skills which enable them to do many things with a fair understanding that they will succeed, barring some extraordinary events like an earthquake or a seizure. Of course, those who accepted Williams's terminology<sup>18</sup> were quick to warn readers that they use 'luck' in a stipulative meaning, to indicate simply what is beyond one's exclusive control. But the choice of terminology is significant. It betrays a willingness to imagine our being in the world as being in an alien environment, where we are tossed about on the waves of fortune any time we venture beyond our inner self.

We need (an outline of) an understanding of our engagement with the world which (a) distinguishes between the ways we gamble, deliberately taking risks, and

---

a certain response, say amused laughter, may be justified, say in response to a joke, even though the fact that one is amused by such a joke shows that one has a rather crude taste. The justification of actions and attitudes is, in other words, limited to them, and is consistent with the fact that they may betray unpleasing, or worse, character traits or dispositions (showing the agent to be suspicious, tight fisted, etc.).

<sup>18</sup> which he introduced with tongue in cheek, intending it to be self-undermining, thus leading to the rejection of the 'morality system' (see his postscript)

the ways our actions, while depending on matters over which we have little influence, are not gambles, and (b) makes plain the roles of these different forms of engagement in the constitution of ourselves.

The distinction between deliberate risk-taking and other actions and activities is inevitably a soft one. We are often aware of a risk, however small. But the distinction is important in demarcating two distinct attitudes, with many intermediate ones, combining features of both. At one extreme are gambling<sup>19</sup> and other actions over whose outcome we have hardly any influence, and where we do not have warranted beliefs about their outcome (except, sometimes, about the chances of their turning out as we would prefer them to do). I mean to include in this category not only playing roulette or betting on horse races, but also other more ordinary enterprises, for example, setting out on a journey by hitch-hiking, not knowing whether any driver will offer us a lift (I assume, perhaps contrary to fact, that little skill is involved in hiking, beyond choosing the spot where one waits).

Other activities are very different. We expect their outcome to depend on our skill and effort. That does not mean that we are not aware that they too depend on factors over which we have little influence, but we believe that we know how they are going to, or are very likely to, turn out. When one walks along a narrow canyon one is aware that a flash flood may put an end to it all, but normally one knows that it is a very remote possibility. When one starts a degree course, or takes employment, one expects to be able to complete one's studies or perform one's duties to one's employer, using one's wits, but aware that various unexpected events may frustrate one, etc. Many activities (cooking, eating, shopping, going to the theatre, visiting friends, etc.) fall into this category.

Many others are mixed cases. In opening a grocery store or a restaurant, I count on my skill to make a success of it, but am aware that an economic downturn,

---

<sup>19</sup> One gambles when taking the risk is an end in itself or a means for the end of gaining whatever is the prize for winning. In other cases in this category, the risk is not the means to the end, but merely a feature of the situation one puts up with. Gambling does not pose the problem for the Control Principle that other risk-dependent outcomes do. Either one is not responsible for the outcome at all, or, in cases where one is, one knowingly undertook the risk of that outcome, and in so doing as it were consented to the outcome.

or the appearance of competition, are not unlikely. One is both taking a gamble, a risk, and relying on one's resolution and skill to navigate one to success.

These two ways in which our activities depend on risk are important in our lives in different ways. The second, normal, case, the case in which one relies on nature to play along (even though one is usually aware that it may not) is crucial to our ability to act (with a modicum of success) at all. To do so we must learn to assess what is likely or unlikely to happen in the normal course of events, to judge whether our situation is normal, and to develop skills which assure us of success in the normal case, by testing our skills to their limits. That is how we learn when we can trust our skills, and depend on nature (including other agents) co-operating in our efforts, and when we are taking risks, placing ourselves at the mercy of luck.

Unless I can trust the chair to carry my weight, the ground not to give way when I move across it, the plate on the table not to be stuck to the table when I reach for it, and to maintain rigidity and balance when I hold it, and so on, I cannot perform even the simplest act. More complex acts require similar though more complex understanding of one's environment. Furthermore, learning how to perform actions mostly involves trying to perform them and developing and honing a skill to do so, gradually testing the limits of one's abilities and skills as one expands them. Failure is in such cases an essential part of the learning process, a process which for the most part is not separate from normal acting. True, at an early age one learns rather than acts, and later on one may undertake periods of training where acquiring the skill is the purpose of the activity. But for the most part, learning is concurrent with acting. The pianist improves as he practices, gives recitals, makes recordings, and so on. Failure remains, throughout our life, part of learning, solidifying, and reassuring one about one's skills and their limits.

That is in large part the way we make ourselves into who we are. Who we are, in the relevant sense, is determined by our dispositions and attitudes which incline us to pursue some goals and keep clear of others. And these dispositions are shaped in large part by our skills, and our awareness of them. They are shaped by our self-image as people aware of their abilities, and willing or unwilling to challenge their limitations, to run or to avoid certain risks, and so on. There are two sides to

that self-image: On the one side, we acquire not only an understanding of many aspects of our environment, and an ability to predict them, but also skills which enable us to make use of our understanding. We acquire and are aware of having a sphere of secure competence, consisting of a range of actions which, in normal circumstances, we reliably expect that we will successfully perform if we set out to perform them, barring competence-defeating events (which are very rare). The range of such actions and circumstances varies between people, and changes over time regarding each one of us. It is marked by the fact that regarding these actions when considering whether to perform them we need not deliberate about the chances of success. On the other side, we develop risk-taking, or risk-averse dispositions, which incline us to take greater risks regarding some matters than regarding others. Both our sense of our secure competence, and our willingness to challenge our limits or to stay within them, are central to our sense of who we are.

Let me summarise: First, our life, its successes, failures and its meaning, are bound up with our interaction with the world, with our impact on it and its impact on us. Second, while in some of our activities we put ourselves at the mercy of luck, and sometimes that may be the point, the thrill, of the activity, in others we rely on our acquired skills, confident, to various degrees, that we know how to succeed given normal natural conditions. The distinction crucial for my case is the distinction between attitudes. They are attitudes to phenomena, the likelihood of various events, which on some accounts present a continuum with no natural breaks. Our attitudes, however, tend to display categorical distinctions even regarding phenomena which can be analysed as continuous. Third, our sense of who we are while in part determined independently of our activities (say by gender or ethnicity and their social meanings) is in part determined by our sense of our abilities and their limitations (always against the background of the natural and social environment of our life), which (in ways dependent on our temperament and dispositions) fixes the limits of our ambitions and aspirations. Fourth, that sense of who we are, though sometimes dramatically influenced by dramatic events, is continuously being moulded through our understanding of our actions, which reinforce, extend or undermine our confidence in our abilities and capacities. Fifth,

the process of shaping who we are that I described is normatively driven, that is we form views of who or what we want to be in light of views of what people, or people like us, should be.<sup>20</sup> Sixth, in that process our actions and their success both reveal who we are and make us who we are, in ways which are often difficult to disentangle.

#### 4. RESPONSIBILITY AND THE GUIDANCE PRINCIPLE

I began by outlining a conception of responsibility<sub>2</sub>, which was only partly justified. It was motivated by showing that it relies on a relationship between what we are responsible for and our capacities of rational agency. But I did not explain why that relationship, rather than some alternative one, determines responsibility<sub>2</sub>. The preceding reflections on our connectedness to the world provide the material for the missing explanation. It comes in two parts. In the present section I will explain that rational guidance reaches further than is often appreciated. In the next section I will explain why responsibility extends to certain cases of failed guidance.

The Guidance Principle, an element of the conception of responsibility I endorsed, holds people responsible both for their independent intentions (i.e. those one can have even while not doing what one intends to do – future-directed, they are often called) and for completed intentional and controlled actions. Both are guided by us through the use of our powers of rational agency, but differently: independent intentions are formed in light of our view of our situation, and of the proper response to it, the case for planning for the future, etc., all manifestations of our powers of rational agency. Intentional actions are typically governed by embedded intentions through which we guide and control them. But note, that it is the action itself, not merely the intention, which is guided by us through our powers of agency.

Some writers on moral luck miss the point. They think that even though we can be said to be responsible for completed intentional actions this is so merely because we are responsible for the intention which produced the action. Hence according to them the consequences of responsibility, the attitudes and responses

---

<sup>20</sup> See Section 5 for further elaboration of this point.

sanctioned by our responsibility for the action, are the same as the attitudes and responses sanctioned by the intention (i.e. the independent intention) to perform it. This seems to be a mistake due to a misguided notion of control.<sup>21</sup> Even aspects of intentional conduct which depend on matters beyond agents' exclusive control are typically guided and controlled by them as I have explained earlier.

Some writers believe that people are responsible for their actions simply because they are responsible for attempting those actions. The completion of the action, i.e. those parts of it beyond the attempt to perform it, are – they maintain – not controlled by the agents who are therefore not responsible for the actions beyond the attempt stage. And, to be sure, in many cases the embedded intention in an attempt to  $\Phi$  is the same as in  $\Phi$ -ing. For example, when one attempts to kill someone by poisoning him, and does kill him by that poisoning. Therefore to the extent that responsibility depends on being guided by the embedded intention it would appear that the agent is responsible for the action because he is responsible for the attempt. Talk of responsibility for the action is mere *façon de parler*.

But the argument is flawed. In typical intentional actions we guide and control not merely the intention but the act, including its result. To be sure, sometimes people intentionally succeed due to sheer luck. A hopeless shot can intentionally hit the bull's eye. But such cases are atypical. Typically, we control intentional acts, including their aspects which depend on factors beyond our control, through their embedded intentions, which guide our movements, adjust them to the circumstances in a way calculated to secure the intended result.

That is also why we are responsible for most attempts. For the most part an attempt to perform one action is the performance of another action. For example, a failed attempt to murder someone may consist in firing a gun and missing the intended victim. So for the most part responsibility for attempts presupposes

---

<sup>21</sup> There can be no denying that one can stipulatively define 'control' to make it true by definition that one can control one's intentions and nothing else. But stipulations do not generate philosophical puzzles. The puzzle of moral luck arises only if we are to blame for actions or consequences which are beyond our control (given the meaning of 'control' in English). Hence my claim that some writers misunderstood it.

responsibility for completed actions because most attempts are completed actions.<sup>22</sup> When assessing an intentional action (e.g. a murder) we assess something quite distinct from what is assessed when assessing an independent intention (e.g. an intention to murder someone). This makes it possible to maintain that the responses and attitudes appropriate to intentional actions differ from those appropriate to an independent intention, consistent with the denial of moral luck.

What of cases in which we are subject to risk, and to luck, as when we bet on the horses, or make speculative investments? It is easy to misperceive the role of intention in such actions. To be sure I intentionally go to the casino, and I intentionally place a bet. But I do not intentionally win the bet, rather, luckily I win. I intentionally make the speculative investment, and I may even intend to make my fortune by making such investment. But I do not intentionally make a fortune through that risky investment. It just happens to turn up trumps for me, as I hoped that it would. To be sure the divide is anything but sharp. The more skill and foresight goes into the action the more appropriate it is to say that I intended its result.

Given that conforming to the Guidance Principle is not necessary for responsibility, the absence of anything like a sharp boundary does not matter. At the present stage of the argument all I am claiming is that quite often when acting intentionally we are in control of the result. Those tempted to reject this conclusion have to deny that we are responsible for most attempts as well, for they too are actions.

What could motivate such rejection? One thought is that the action successfully completed on this occasion might have failed had factors beyond our control intervened. But that can only establish that had we failed because of such

---

<sup>22</sup> The exceptions are those special circumstances in which one can try to do something (normally a basic action) without performing any action, as when one regains consciousness after an accident and tries unsuccessfully to move one's arm (such tryings are, like some other doings, not themselves actions). Normally, we cannot try to move our arm, though we can move it. These controversial claims have been much discussed. I do not rely on them in any way. When we do attempt a basic action, we are responsible for the attempt, and sometimes we are responsible for not making the attempting. My only claim relevant to the purpose of this paper is that we are also responsible for other attempts, like attempted murder, which are completed acts.

factors our failure would have been beyond our control, and we are not – if the Guidance Principle sets a necessary condition for responsibility – responsible for the failure. That does not establish that in the circumstances which actually existed at the time of action we were not in control of the action and its result.

Another thought is that our successfully performing the action depended not only on our intention or attempt but also on other factors (we succeeded in getting from street to kitchen because the lock functioned well, the floor supported our weight, etc.). These factors, the thought is, are not under our exclusive control, and therefore the action is not under our exclusive guidance and control. But that last step is a non-sequitur. The action is under our guidance and control because we could and did adjust our action to the prevailing circumstances, took advantage of them, avoided difficulties they presented, etc. That is the way control is exercised when we engage with the world.<sup>23</sup>

In his 'Involuntary Sins'<sup>24</sup> Robert Adams reminds us, putting his point in the terms of the present discussion, that people are often held responsible for their emotions (e.g. for excessive, irrational anger, or for jealousy), for their beliefs (e.g. that some races are inferior to others), and for other attitudes (e.g. self-righteousness) and are thought to be blameworthy (for beliefs, emotions and attitudes like the above) or praiseworthy for them.

Adams explains why we are responsible for some of our psychological states:

*The deepest reason for accepting this responsibility ... is that it is rightly ours. It is important for a correct ethical appreciation of one's own life. To refuse to take responsibility for one's emotions and motives is to be inappropriately alienated from one's own emotional and appetitive faculties. (16)*

That is true, but – as Adams recognizes – if unqualified it goes too far. We also feel cold when temperatures drop, disoriented when our blood pressure drops, but we

---

<sup>23</sup> Zimmerman ('Luck & Moral Responsibility' in D. Statman (ed.) *Moral Luck*, 1993) rightly warned against understanding "control" in the Control Principle as some kind of exclusive control on all the conditions for the occurrence of what is supposed to be under control.

<sup>24</sup> R.M. Adams, 'Involuntary Sins' *The Philosophical Review* 94 (1985) 3-31. I believe that most of the sins Adams deals with are neither voluntary nor involuntary. They are psychological phenomena to which the distinction does not happily apply. See for more detailed discrimination J. Raz, *Engaging Reason* pp. 11-12. Adams identifies the voluntary with what is chosen or meant. That seems to me inaccurate. See above pages 12-14.

are not responsible for these psychological phenomena (though sometimes we are responsible for their causes or their consequences). Therefore Adams qualifies the explanation, restricting it to a limited class of psychological phenomena:

*My suggestion is that among states of mind that have intentional objects, the ones for which we are directly responsible are those in which we are responding, consciously or unconsciously, to data that are rich enough to permit a fairly adequate ethical appreciation of the state's intentional object ... (26)*

I think that in that explanation he is right. Moreover, the explanation shows how responsibility for beliefs, as well as for some feelings, emotions, and other attitudes, is inherently inconsistent both with the Control and with the Guidance Principle. I have argued before for the fundamental importance of the distinction between the active and the passive aspects of people's lives, with the active being those aspects of our life in which we respond to reasons (practical or adaptive) as we see them.<sup>25</sup> We control our beliefs, I suggested, and are in control of our emotions, desires, intentions and actions in so far as we respond to reasons as we see them, and have those beliefs, emotions, desires and intentions which we take to be in accord with reason. We are not in control; we are tossed hither and thither, when in the grip of urges, passions, moods or emotions which we take to be out of line with reason.

##### 5. BEYOND THE GUIDANCE PRINCIPLE

To remind ourselves: I am looking to explain why conduct for which we are non-derivatively responsible<sub>2</sub> is related to powers of rational agency in accord with the **Rational Functioning Principle**, namely that we are non-derivatively responsible for conduct which is the result of the functioning, successful or failed, of our powers of rational agency. As we saw, even the **Guidance Principle**, the relatively uncontroversial element of the **Rational Functioning Principle**, establishes responsibility<sub>2</sub> for some aspects of conduct over which we do not have exclusive control. I will assume that its connection to our powers of agency requires little more than further elaboration of its details and implications. Not a task for this paper. What is needed here is an explanation of those aspects of the **Rational Functioning Principle** which reach beyond the **Guidance Principle**, namely those which assert

---

<sup>25</sup> *Engaging Reason* ch. 1, see also David Owens, Susan Hurley

responsibility for acts and omissions due to the mal-functioning of our powers of rational agency. It is here that we find the examples of “moral luck”, which many found troubling.

The Rational Functioning Principle applies only to people who have powers of rational agency, and only when these powers are not blocked or suspended (as they are by sleep, heavy sedation, and the like). It asserts that people are responsible for conduct which is due to their powers of rational agency. People vary in the extent of their powers. Their abilities to absorb information, to use it in deliberation, to reach conclusions and be moved to conduct accordingly, as well as their ability to control their bodies, the range of actions they can perform with confidence, and the circumstances in which they can do so – vary considerably. The principle takes this into account. It asserts responsibility<sub>2</sub> for conduct due to the powers of rational agency which the person in question has – the range of conduct one is responsible<sub>2</sub> for may therefore vary according to the range of one’s powers of agency at the time.<sup>26</sup>

Thus the principle connects with my earlier observations regarding our domain of secure competence. We are non-derivatively responsible<sub>2</sub> for unintentional actions only if they are the results of a failed intentional action which falls within our domain of secure competence. Only then is the action due to a failure of our powers of rational agency, in the meaning of the principle. The second way in which the Rational Functioning Principle goes beyond the Guidance Principle is in affirming our responsibility for unintentional omissions due to failure of our powers of agency, though in these cases the failure is mostly of our mental powers. Failures of both kinds are normally classified as negligence.<sup>27</sup>

---

<sup>26</sup> See more on this in my “Being in the World: The Limits of Responsibility” also available on my personal website. It does not mean, of course, that people who stray beyond their spheres of secure competence are not responsible<sub>2</sub> for doing so. They take risks and (by the combination of the principles of basic and derivative responsibility) may well be responsible<sub>2</sub> for taking those risks.

<sup>27</sup> Negligence includes a normative element: failure of performance of a certain kind which **should not have occurred**. Given that the cases I am concerned with are ones in which one failed in doing what one intended to do, that one ought not have failed follows from the assumption that one’s intentions were reasonable. It would follow, by the facilitative principle, that one had reason to execute them successfully.

Acts within the domain of secure competence are ones we are entitled to undertake without reflecting on the prospects of successfully performing the action (though the exemption does not extend to the case for its performance). Agency presupposes the availability of such actions. I have suggested that they are central to our sense of who we are, our sense of our own identity. In acknowledging responsibility<sub>2</sub> for actions due to our rational powers we are simply affirming that they are our secure rational powers. Our sense of who we are, which underpins our self-esteem, as well as our inclination to take or avoid risks and therefore our aspirations and ambitions, is tied up with our success in establishing a domain within which our powers of rational agency are securely reliable. In holding ourselves competent within that domain we hold ourselves responsible<sub>2</sub> for actions which fall within it. To disavow responsibility<sub>2</sub> for such actions is to be false to who we are.<sup>28</sup>

Must it be so? Could it not be otherwise? Those who look for an argument from first principles will be quick to point out that I provided no such argument. I believe that none is possible. The aim was to point to the significance of correct assignments of responsibility<sub>2</sub> in our life and thought. The above is part of discharging the task.

## 6. BUT IS IT BLAME?

Williams drew attention to the fact that the consequences of our actions matter to us even when they are due to luck. Critics remarked that having raised the puzzle of moral luck his discussion fails to address it. Agreeing with both I singled out one type of consequence which matters to us even when it is neither intended nor foreseen. Consequences matter to us because we are responsible<sub>2</sub> for their occurrence.

Responsibility<sub>2</sub> marks a phenomenon which, though manifesting the significance to our life of consequences of our conduct which reach beyond our exclusive control, is narrower than agent-regret, and yet is wider than being worthy of blame or praise. There is a whole variety of responses and attitudes mandated by responsibility<sub>2</sub>, other than blaming or praising. At the practical end of the range

---

<sup>28</sup> Adams's observation quoted earlier though confined to our emotions applies to actions as well: "to refuse to take responsibility for one's emotions and motives is to be inappropriately alienated from one's own emotional and appetitive faculties."

stands the duty to make amends when one wronged someone by conduct for which one is responsible<sub>2</sub>. The availability of excuses, those which do not negate responsibility<sub>2</sub>, does not exempt one from that.

The difficulty which some find with the puzzle of moral luck may nevertheless trouble even those who accept the **Rational Functioning Principle**. Focussing on blameworthiness, they are perplexed by the tendency to judge the negligent killer more blameworthy than the negligent driver whose good luck it was not to harm anyone. As Nagel put it: 'one cannot be more culpable or estimable for anything than one is for that fraction of it which is under one's control' (Nagel, 28).

Williams was indifferent to that worry. Undermining "the morality system", or at least bringing out its limitations, was one of his goals. By and large my reaction is similar, though I am unable to explore the notion of blame here. One thing is clear, to undermine successfully the misguided notion of blame it is essential that the notion of blameworthiness involved in cases which are inconsistent with the denial of moral luck is a familiar and important notion of blameworthiness. I will not say much in support of this contention, for I see little reason to doubt it. Blameworthiness is a broad category, encompassing all wrongdoing for which we are responsible and not excused.<sup>29</sup> It ranges from the trivial to the serious. Furthermore, various invocations of blame will carry different pragmatic implications, and these accommodate all the distinctions which are called for.

Some will agree to the above, but insist that they have in mind a particular kind of blame only: moral blame. I do not see a case for assuming that there is here, any more than elsewhere, a systematic and theoretically significant distinction between the moral and the rest.<sup>30</sup> More importantly, blaming can, and often has, a very specific object: I can not blame Jones for his action and yet blame him for the attitude, the intention, the beliefs that his action expresses. Or I can blame him both for his action and for his attitude, and so on. There is no case for maintaining that

---

<sup>29</sup> Some excuses establish that the agent is not responsible for the action. Others are compatible with responsibility but negate blameworthiness.

<sup>30</sup> See on the problem of distinguishing the moral from the non-moral *Engaging Reason*, ch.11

there are here different concepts of blaming. There are simply different objects of blame.<sup>31</sup>

What we needed, and the paper aimed to supply, is an understanding of how it could be that we are responsible and therefore may be to blame for aspects of conduct which exceed our exclusive control, and may be differentially to blame for intentions, attempts, and actions which share their mental component. I will summarise briefly how the argument of the paper helps with this.

Consider intentional actions first, and accidental ones second. Some writers hold that the degree of blameworthiness or praiseworthiness of an intentional action is identical with that of an intention to perform it. Given the understanding of control I suggested earlier this would not be the consequence of Nagel's dictum. When acting intentionally we normally control the results and the foreseen consequences of our actions. But do we deserve praise or blame for them to a degree different than the praise or blame their intentions merit?

Intentions, like emotions such as fear, anger, pride, gratitude, jealousy, which also express ways of psychologically relating to and reacting to the world, are subject to judgements of responsibility, blameworthiness etc. But as they are part of our interior lives, and merely aim at action, there is a good case for judging them on their own, and for not applying to them automatically the judgements which would be appropriate to apply to the actions or omissions they led to. That is the lesson of the discussion of the meaning actions and their consequences have for us. The lesson has to be applied with attention to different cases, but I will leave this matter here without exploring these implications.

Perhaps, while actions may merit blame and praise in ways which differ from the independent intention to perform them, attempts to  $\phi$  necessarily merit the same praise or blame as  $\phi$ ing. That would tend to suggest not, as most writers do, that we are too lenient in our views of the degree of blame attached to failed attempts, but that we are too severe in the degree of blame we think that successful

---

<sup>31</sup> There is a case for thinking that blaming the wind, just like believing that the wind is responsible, for the damage is blaming in a different sense from blaming a person for an unexcused wrongful action for which he or she is responsible.

wrongdoers deserve. After all, we judge successful wrongdoers more severely because they succeeded, but their success – according to this argument – does not add to the blameworthiness.

Of course, the fact that intentional actions as a whole, and not only their intentionality, are under our control only means that the Rational Functioning Principle does not rule out the possibility that the degree of blame or praise their performance earns is affected by their results. Is there any reason to think that they actually have such an effect? We are told that most people believe that they do, for it is a widespread view that one is to blame more for a wrong done than for a wrong attempted. Williams's discussion, and my elaboration on his theme, explained why that belief makes sense. They explained why we care about the results of our actions, and often judge ourselves by our success or failure to achieve them. They explain the significance of the attachment to our actions, results included, and of the fact that we judge ourselves by our ability to achieve them in our life. This is consistent with the fact that agents who completed an intentional action may be to blame both for the action and for intending it, whereas if they tried and failed they are not to blame for the action (which did not take place) but may be to blame for the intention, and that blameworthiness is the same as the blameworthiness of those who complete the action for their intention to perform it (assuming that the circumstances are the same).

In one respect matters appear different when we deal with accidental actions. Think of an action (the bringing about of a result) which one regrets. If it was intentional the agent embraced that regretted result, either accepting it as a foreseen consequence or aiming to achieve it. This "accepting" connects the agent to the regretted action and its consequences. That is why the agent may be to blame for them. If the action was merely negligent, and therefore accidental, the agent may not have foreseen the result. He regrets it now. But that regret is set aside from his responsibility<sub>2</sub> for the action. Think of the driver who negligently kills. Of course he regrets the killing. But he would regret it also had he not been negligent, and the killing were a mere accident. His regret is different because he was negligent. Crudely perhaps we will allow that it is greater in that case. But that may be because

there are two things he regrets: his negligence and his killing. But they are unconnected. His regret for the killing is the same regret he would have felt had he killed without being negligent. Therefore, some would say, the agent cannot be more to blame for the negligent killing than for the negligent driving.

What I suspect we actually do, and certainly what we should do, is to refuse to judge all negligent agents with the same brush, one coloured by the facts which make them negligent and nothing more. Quite apart from the significance of their conduct and its consequences, their mental states differ a good deal, both on the occasion and as reflecting a habit of mind and action. In some cases of negligence agents actually foresee the accident they bring about. In some others they foresee the probability, or the likelihood, or the possibility, of the accident, with a greater or lesser degree of clarity and of determinacy. And sometimes they welcome it, at others they are indifferent to it, at others still they hope to avoid it, and again these appear in various degrees of clarity and determinacy. To claim that because they would have regretted the outcome even had they not been responsible for it, and because their responsibility is for negligence only they have failed to “accept” the outcome, and their regret is independent of their responsibility, is to be fixated on a thin distinction among mental attitudes which bears little relations to people’s experience, and to their normative understanding of themselves and others.

Those who think so could allow that many other reactions may be warranted by negligent killing and not by negligent driving which harms no one. But, they insist, when it comes to blame (or praise) evaluation is determined by the agent’s attitude (or the part of his conduct which he controlled), and that evaluation is central to moral standing.

The previous analysis helps expose the artificiality of this conception of blame. There is no reason to deny that it is possible to form an attitude to negligent agents (and to others) which conforms to that artificial conception. Nor is there a need to deny that such an attitude has a reasonable role to play in people’s attitudes to themselves and to others. But to assume that it has some unique moral standing, even though it ignores everything else, including the importance of the object of the agents’ intentions and beliefs, the significance of their actions, and all more nuanced

distinctions regarding their mental states, anything other than what they intended or knew, is the mark of living in a blinkered, impoverished and distorted world.

## **Appendix: More On Williams's account**

### **A. CASTING OFF THE SUBJECTIVIST ELEMENTS IN WILLIAMS'S ACCOUNT**

In responding to Williams Nagel points out that not only did Williams say nothing about moral judgements. He said nothing about any objective judgements, having confined his discussion to a feeling that the agent himself need not stand by.<sup>32</sup> It has to be conceded that Williams goes out of his way to underline the fact that *Gauguin's* retrospective justification is not one which his abandoned family is morally bound to accept. But that may be more the result of Williams's rejection of moral realism, than a product of his view on moral luck. I will continue to explore the issue presupposing some form of realism about reasons, and therefore about justifications.

To the extent that emotions and attitudes can be justified or unjustified so can emotions of pride, agent-regret, shame and their like. Williams correctly observes that as one forms oneself through one's biography various events and emotions may change their meaning. This shift in meaning is inevitable given the fact that their meaning changes with one's actions, and with other events in one's life, events which are beyond one's control. But that does not undermine the possibility of true beliefs about those meanings, and about the justification of the agent's actions in light of those meanings. It merely means that these judgements are time-indexed. What is true today may not have been true a month ago. That, rather than the rejection of objectivity, is the lesson of Williams's point about retrospective justification, and changes of meaning depending on the outcomes of actions. Given

---

<sup>32</sup> 'Williams sidesteps the fascinating question raised in his paper. He does not defend the possibility of moral luck against Kantian doubts, but instead re-describes the case which seems to be his strongest candidate in terms which have nothing to do with moral judgment. Gauguin's talent as a painter may be a matter of luck, but it does not, according to Williams, warrant the retrospective judgment that his desertion of his family is morally acceptable. In fact, it does not warrant any judgment about his prior decision that pretends to objective validity for everyone, or even to timeless validity for him.' (137)

that changes of meaning often take time and are not due to a momentary act or event, we can also expect that at times the meaning and success of acts and projects will be indeterminate.

The very possibility of *Gauguin's* action being justified from his point of view means that it can be unjustified as well, and that *Gauguin* may be mistaken about whether it is justified. His belief that it is justified does not make it justified. That much is part of the very possibility of justification (special cases apart) and is implied by Williams's own discussion.

The distinction between good or valid justifications and spurious or mistaken ones can be known by people other than the person whose action is in question. That person may have privileged access to some information relevant to the validity of justifications, but that information can in principle be shared with others, and to the extent that the person concerned is in a privileged position to evaluate and assess the facts, such privilege is unlikely to be absolute, and in any case others can when this is appropriate defer to the judgement of the agent. So in a fairly straightforward sense if the justification is valid it is not only valid for the agent. Its validity can be appreciated by others too. This is consistent with those other people believing that the agent should compensate those who were injured or disadvantaged by the action that they agree to be justified. It is also consistent with other people believing that they would not have acted as the agent did, or that a better person would not have acted in that way. Justifications do not establish that the action was required or obligatory, only that it was permissible.

So Williams may have been too reticent when he denied the objectivity of the justifications he was discussing, or their objective, inter-subjective applicability. Sometimes what makes particular justifications valid defies our ability to produce sound generalisations. But there are some generalisations we need to hazard. One challenge which has to be faced is the claim that while agent-regret may sometimes be justified it is never justified when the regret is due to factors beyond the agent's control.

## **B. WILLIAMS'S JUSTIFICATION OF AGENT-REGRET**

If I understand him correctly, Williams's reply to the question about the significance of agent-regret is less successful: it is confined to important decisions. Regarding most decisions agents' concern about the success of their actions in terms of the quality of their consequences, Williams tells us, is overshadowed by their concern for the quality of the deliberative process. If our deliberative processes are faulty they are likely to fail again. The adverse consequences of faulty process may reach far and wide, well beyond the adverse consequences of the current decision. In such cases regret occurs only when we find fault with our deliberative procedures, and that means that we assess our decisions in light of the circumstances at the time they were taken.<sup>33</sup> But there are other cases as well:

*In these cases, the project in the interests of which the decision is made is one with which the agent is identified in such a way that if it succeeds, his standpoint of assessment will be from a life which then derives an important part of its significance for him from that very fact; while if he fails, it can, necessarily, have no such significance in his life. If he succeeds, it cannot be that while welcoming the outcome he more basically regrets the decision; while if he fails, his standpoint will be of one for whom the ground project of the decision has proved worthless, and this (under the simplifying assumption that other adequate projects are not generated in the process) must leave him with the most basic regrets. So if he fails, his most basic regrets will attach to his decision, and if he succeeds, they cannot. That is the sense in which his decision can be justified, for him, by success. (132)*

Agent-regret is justified<sup>34</sup>, Williams seems to be saying, in cases in which 'the projects' for the sake of which the actions were taken are of a sort that endow an agent's life with some of its meaning. In such cases the outcome matters in itself and justified agent-regret will turn on the success or failure of those projects.

---

<sup>33</sup> Williams confines agent-regret to regret for what one does. Does this extend to taking decisions and forming intentions? Forming an intention is not an action, but if for that reason agent-regret does not apply then there is no reason to apply it to making a decision, which is an action, albeit a mental one. We can either extend the application of agent-regret to the having of attitudes like intentions, and others, or exclude mental actions from its ambit. I will follow the second course here. This means that agent-regret always relates to an occurrence which involves an element of luck, for the successful completion of an action (other than a mental action) always depends on elements beyond the control of the agent.

<sup>34</sup> I will continue to interpret his discussion as applying to justification as well as to rationality.