

2001

The Moral of Moral Luck

Susan Wolf
Johns Hopkins University

Follow this and additional works at: http://digitalcommons.brockport.edu/phil_ex

 Part of the [Ethics and Political Philosophy Commons](#)

Repository Citation

Wolf, Susan (2001) "The Moral of Moral Luck," *Philosophic Exchange*: Vol. 31 : No. 1 , Article 1.
Available at: http://digitalcommons.brockport.edu/phil_ex/vol31/iss1/1

This Article is brought to you for free and open access by Digital Commons @Brockport. It has been accepted for inclusion in Philosophic Exchange by an authorized editor of Digital Commons @Brockport. For more information, please contact kmyers@brockport.edu.



The College at
BROCKPORT
STATE UNIVERSITY OF NEW YORK



Susan Wolf

The Moral of Moral Luck

Susan Wolf

In 1976, Bernard Williams coined the phrase “moral luck” to refer to the range of phenomena in which our moral status - how good or bad we are, and how much praise or blame we deserve - is significantly determined by factors beyond our control.¹ Whether we are naturally sociable or irritable, whether we find ourselves faced with particularly explicit or burdensome moral challenges, whether the arrows of our actions hit their targets, all constitute ways in which things we cannot control affect the moral quality of our lives. All, then serve as examples of “moral luck,” which, taken as a group, comprise one of the most philosophically perplexing and troubling features of ordinary moral experience.

To accept the phenomena uncritically is to allow that one can be praised or blamed for what one cannot help. This goes against a very deep commitment most of us have to the idea that you should be morally judged only for what falls within the sphere of your will. Yet to reject the judgments and practices that seem unavoidably to lead to these phenomena would require a radical, and perhaps practically impossible, revision of ordinary moral evaluation.

In this essay I shall be concerned primarily with one type of moral luck, luck in “how things turn out.” A paradigmatic case is that of the truck driver (or, in Williams’ essay, “a lorry driver”) who accidentally runs over a child. Let us assume that the driver is guilty of a minor degree of negligence - he has not had his brakes inspected as recently as he ought - and that this negligence contributes to the accident. What makes this a case of moral luck, if it is a case, is that this truck driver has much more about which to feel guilty - he has much more moral weight on his shoulders, so to speak - than other drivers who, though equally negligent, had no children run across their paths.

I shall discuss the truck driver example, and variations of it, at great length in this essay, and shall occasionally refer to one or two other instances of moral luck. It should be noted at the outset, however, that the phenomenon in question is ubiquitous. Every day, people in laboratories, government offices, corporations, and universities sign off on things to which they ought to put a stop, or bend the rules for the sake of convenience or laziness or misplaced generosity. Rarely, but occasionally, such acts of flawed reasoning, negligence or weakness blow up in someone’s face. What is philosophically in question is how we should judge the perpetrators of these acts, and how the perpetrators should judge themselves. Specifically, the question is whether those whose acts actually lead to serious harm deserve the same treatment and the same judgments as those who, but for fortune, would have caused as much damage.

The problem of moral luck

In the discussions of this issue I have witnessed, only a few people (includ-

ing perhaps Williams himself) seem to accept the idea of moral luck more or less wholesale. That is, only a few seem to hold in a perfectly unqualified way that a person whose actions have morally worse effects is herself worse or more blameworthy than one whose equally faulty behavior has less harmful consequences. The majority tends to reject the idea that there really is such a thing as moral luck, drawing instead the conclusion that many of the judgments we make in day-to-day life are simply inconsistent. According to the majority's position, although we may in fact blame the driver whose recklessness causes a death more than we blame the equally reckless driver who causes no harm, the blame they deserve is equal. We would be more just and more rational, according to this view, if we were to regard these agents, and if these agents were to regard themselves, as equally faulty and equally blameworthy.

The latter view, according to which equal recklessness deserves equal blame, I shall call the rationalist position. To mark the opposition, I shall call the former view the irrationalist position. These labels, however, should not keep anyone from assessing these positions on their merits.

I myself have some sympathy, if also some dissatisfaction, with each position. The rationalist view, that equal recklessness deserves equal blame, seems to me, on first reflection, the more justifiable. Yet the ideal of justice and rationality that would have us regard two agents whose actions have had drastically different moral consequences as being in precisely the same moral position, and especially, the ideal that would have the two agents regard themselves as being in the same position, strikes me as not just unrealistic but positively eerie.

Something, it seems to me, is wrong with the rationalist position - or perhaps I should say that something is missing, for, as shall become clear, I do not believe that the rationalist position is, strictly speaking, false. Rather, it is incomplete. It fails to capture, or even to acknowledge, something morally significant in the phenomenon of moral luck, something that suggests some grain of truth in the irrationalist response. My goal in this paper is to bring out more clearly what is unsatisfying about the rationalist position, and to find an acceptable way of expressing the grain of truth in the irrationalist response. In short, I hope to find the moral in the phenomenon of moral luck.

The rationalist position considered

Let us begin by looking more closely at the rationalist position, according to which luck in how things turn out is irrelevant to blameworthiness. According to this position, how much blame one deserves depends on how wrongly one has acted, on how much moral fault is revealed or expressed or instantiated in the action one has performed (or in one's failure to act). Of course, assessing how wrong or morally faulty an action (or failure to act) is is itself a highly complex matter,² but one thing moral faultiness is not a function of, according to this position, is how the action turns out. The robber who tries to shoot a storekeeper in the heart is as blameworthy as the robber who succeeds, at least if the only difference is that the former's pistol misfired. The parent who breaks his child's jaw in a fit of rage is as blameworthy as the parent who kills

his child, if the only difference lies in the fortuitous angle of the child's fall.

This position, as I have already noted, seems by far the most natural one. Why should we even hesitate to accept it? To be sure, our actual judgments and attitudes do not always conform to rationalist principles. But, a defender of rationalism will suggest, these data can be explained without thereby being justified. Indeed, when we consider the circumstances in which such judgments and attitudes are formed, several possible and plausible explanations suggest themselves.

Most obviously, we are more apt to know of someone's attempt at murder, assault, or theft if her attempt has been successful, and we are more apt to know of someone's reckless behavior or negligence if the act or omission resulted in the harm the potential for which made the behavior reckless or negligent in the first place. Since we can only blame people for actions or omissions of which we are aware, we will more often blame people whose wrongful actions or faulty behavior lead to bad results than those whose actions cause no harm.

Second, many if not most morally faulty actions are such that it is difficult to assess how morally faulty they are. Faultiness is at least partly a function of how great a risk of harm, and how great a harm one's actions can be expected to incur. But we often lack the information that would allow us to fix on even an approximate range of expected harm. How reckless was the truck driver's driving? How negligent was the doctor who failed to diagnose her patient's rare disease? Since we cannot be maximally careful, thoughtful, sensitive, protective of others all the time, and since it is hard to weigh the costs of uncertain risks against the benefits, the question of when and how much our actions fall below the line of our duty is not easy. When damage is done - a child is run over, a patient falls into a coma - we tend to suspect that the behavior that led to it was faulty. Chances are, the driver was driving recklessly, the doctor was cutting corners. When things turn out all right, we are less certain whether and how much blame was deserved.

Third, whether or how much we actually blame people varies not only with how blameworthy we think they are (how much blame, in other words, we think is deserved and appropriate), but also with how much we are inclined or motivated to "go in for" blaming. If an act results in harm, we are not only more likely to be aware of its faultiness, we are also more likely to be upset or angered by it. The impulse to blame often provides victims and those who sympathize with them with a way to direct the emotional energy aroused by their suffering. If a building that was not up to the standards of the fire code catches fire and burns to the ground with tenants trapped inside, we are, of course, more apt to find out that the landlady and the inspector she bribed have acted wrongly. But our blame might also be harsher than our blame in another case in which a similar violation was caught before any damage actually occurred, for the urge to blame may be greater as a response to the greater shock, anger, and sorrow at the victims' actual losses.

These and other differences in our epistemological and motivational positions with respect to acts with morally important differences in outcome go a

long way in explaining our differential reactions to these acts. Importantly, they help to explain the reactions not only of the victims of these acts and of third parties who hear of them; they also bear on the reactions of the agents themselves. The question, however, is whether they explain enough to put the matter of how to respond to differential harm to rest. Can our tendency to blame people differentially on the basis of harm caused rather than fault exhibited be completely explained by appeal to our imperfect knowledge and our nonrational emotions?

If so, it seems that we should accept the rationalist position in its entirety, and conclude that, at least under conditions of more perfect knowledge, it would be more just, and simply morally better to blame people equally if they exhibit equal fault, and, indeed, that it would be more just and simply morally better if people whose behavior was equally faulty blamed themselves equally. If we reflect more on specific cases, however, even cases in which the epistemological problems are absent and in which our emotional reactions are subdued, we find some reason to resist this conclusion.

A problem with the rationalist position

Let us return to our two hypothetical truck drivers, this time explicitly assuming the epistemological problems away, by stating that the drivers had possessed equal reason to get their brakes checked and had equally neglected it, with the result that the risk of accident they were incurring was ten percent higher than it would otherwise have been. In the case of the one driver, a child runs into the street and is killed. The other reaches his destination without incident. The position we are considering says that the two drivers are equally blameworthy for driving under these conditions, whatever the difference in the consequences of their behavior.

The suspicion that there is a problem with this position emerges when we follow it up with what seems to be the natural question: The two drivers, we are told, are equally blameworthy - but how blameworthy is that? We may partially answer this question with further comparative claims: this sort of reckless driving is less wrong, less faulty, than increased degrees of recklessness (not to mention than intentional homicide) and more faulty than other instances of driving under some less impaired conditions. But we also seem to want some noncomparative assessment, some more or less determinate measure of wrongness, on the basis of which we might establish some appropriate degree of punishment, or scolding, or even of anger or resentment to be directed toward the wrongdoer. It is quite unclear, however, how one could go about providing this.

Let us look in particular at the question of what level of blame it would be appropriate for the drivers to assign to themselves. We may imagine that the driver whose recklessness led to the child's death feels terrible about what he did. He knows that he did not mean to kill anyone, and that it was an unlucky coincidence that the child ran out when she did. Still, he keeps mentally confronting the fact that if he had had his brakes checked the week before,

the accident might have been prevented. He feels a need to do something significant to help the family of the child, or at least symbolically to express his guilt and to offer penance. He is plagued by nightmares recalling the crash, and cannot drive again for many weeks.

The second driver, let us imagine, arrives home safely and goes to bed. The next morning he reads a story of the first driver and the accident in the paper. Realizing that, but for fortune, that might have been he, he immediately makes an appointment to fix his brakes, taking the time from his busy schedule that he had not previously found. Moreover, he, like the first driver, has learned his lesson once and for all - never again does he let time slip away before checking or repairing his brakes. Still, he gets into his truck that very day and drives without difficulty - he suffers no nightmares, and feels no need to donate time or money either to the cause of safer driving or to the family of the deceased child.

It seems only natural to summarize these descriptions by saying that the first driver blames himself more than the second. If the simple rationalist position on moral luck is correct, this would imply that at least one of these men is not acting, or at least is not feeling as he ought. According to that position, since these two men are equally blameworthy, they should blame themselves to an equal degree. To what degree? The defender of rationalism has several options: She can say that they both ought to feel as the first driver does feel, that they both ought to feel as the second driver feels, or that the right amount of blame falls somewhere in between. All of these answers seem wrong, however.

In fairness to the rationalist, we should note that she may allow what pre-theoretical reflection suggests - namely that there need be no single, uniquely correct reaction in any case, no precise quantity of blame, anger, or punitive suffering that it would be exclusively proper for the agents to impose on themselves. Rather, rationalists may admit that there is a range of reactions each of the agents might have which would seem normal, acceptable, and healthy. The range, however, is not so broad as to stretch, for either agent, across the entire span I have described. If the first driver, who had actually killed the child, responded as I described the second driver doing, we would, I think, be appalled and condemning. Yet if the second driver reacted as the first driver did, inflicting himself with equal guilt, we would also find this disturbing, indicative perhaps, of a psychic imbalance. These judgments might draw one to the third option - that the right degree of blame, guilt, and so on, lies somewhere in between. If so, however, I suspect that it is the appeal of the rationalist view in the abstract that draws us. Were it not for the fact that the drivers' reactions are so different from each other, there would be nothing internal to their separate reactions that would seem in the least inappropriate or odd.

The irrationalist position considered

If one is moved by these considerations, one may be tempted simply to embrace the opposite position, the position I earlier labeled the irrationalist position. For if one thinks it clearly appropriate that the driver who killed the

child blame himself more than the driver who didn't, mustn't this be because he deserves more blame? Following this train of thought, there is at least one point that fairly leaps out in support and explanation of it - namely that, the first driver did something worse than the second: he ran over a child. According to this position, we may suppose, both drivers deserve equal blame for reckless driving. But, in addition, the first driver deserves blame (presumably, a lot of blame) for running over a child. Since the second driver did not run over a child, he cannot deserve blame for that. And so the first driver deserves more blame than the second.

Despite a certain rhetorical strength to this elaboration of the irrationalist position, its force as an argument is superficial at best. It seems a linguistic trick is being used in support of a moral claim. For although it must be granted that the one driver did something that the other driver did not, the explanation for this difference lies wholly in events outside of the drivers' control. Since the issue in question is precisely whether blameworthiness can legitimately depend on differences in outcome that lie beyond the agents' control, an appeal to this difference between the two drivers' actions simply asserts the irrationalist position - it does not justify it.³

Moreover, whatever intuitive support the irrationalist position gains from taking the points of view of the reckless drivers seems to vanish when we turn our attention to the question of how others ought to feel and to judge them.

I earlier suggested that we would expect the truck driver who runs over a child to feel very bad about what he has done, much worse than we would expect or want an equally reckless but much luckier driver to feel. Yet, as the drivers' friends, or even sympathetic observers, it would be appropriate to try to make the first driver feel less bad, to emphasize how limited was his faultiness and how large the element of luck. In other words, it would be appropriate for third parties to bring out the equal faultiness in the two drivers - this, despite our sense that the two drivers themselves ought not, at least to begin with, feel equally bad. The point is not that we expect the first driver to feel worse than we think he ought to feel, and that therefore third parties are required to bring his guilt feelings down to the appropriate level. Rather, it seems, we want the first driver to feel worse - he ought, at least initially, feel so bad that some soothing, some appeals to "reason" are necessary to stop him from judging himself too harshly.⁴

Reconciling our intuitive responses

From the standpoint of impartial observers, the rationalist position seems to assert itself. Yet there seems to be a certain appropriateness to the phenomena of unequal blame in the self-imposed feelings and judgments of the agents themselves. Can we make rational sense of this pair of intuitive responses? Is there any way to reconcile the thought that the equally reckless drivers are equally deserving of blame with the thought that it is nonetheless right for them to feel differently, for the one, as one naturally puts it, to blame himself more than the other?

Since the salient difference between the two drivers is that the one caused a great harm that the other did not, the question is whether it makes sense to believe that one ought to feel bad about causing harm, over and above what one feels for the recognition of one's faultiness in acting in a way that brought about the harm. The issue may seem puzzling because once one's faultiness is factored out, it is unclear why one's special connection to the harm isn't factored out as well. Of course, any decent person will be sorry to learn that a child has been run over, that a woman has fallen into a coma, that a building with tenants inside has burned to the ground. But this will not account for the specific agent-regret, to use Williams' term, that we approvingly imagine the relevant driver, physician, or building inspector to feel. The second driver who reads of the other's accident has both reason to blame himself for his own reckless driving and reason to grieve over the poor child's death - but the feelings these reflections arouse will not add up to the feelings of the first driver whose behavior actually brought about that death. If we cast this puzzle aside, however, and simply accept the judgment that the first driver reasonably feels worse than the second about himself, a different puzzle confronts us - namely, how can it be right for the driver to feel worse about himself and yet not right for us to feel worse toward him?

An answer worth considering is that the unlucky driver has reason to feel badly because things have turned out bad for him, much worse than they turned out for the equally reckless but luckier driver. He has killed a child, and this is something he presumably would very much have preferred to avoid. At the expense of stretching a term, we might say that he has failed in one of his projects. At any rate, we may liken the driver's reaction to the reaction of one who has failed in a project - one who has ruined his marriage, botched an experiment, fumbled the ball. It seems perfectly natural that a person should feel bad about failing to reach his goals. Moreover, it seems perfectly natural that others do not feel badly toward him on these accounts. It is too bad for him if his goals are not met, but he does not deserve blame for it.

There are two problems with the offered analogy. The first is that it assimilates the driver's regret for having killed the child to the case of someone feeling sorry for himself, as if, in addition to the blame he feels for having driven recklessly, he is sorry about his bad luck. "Why did the child have to run across my path, rather than the other guy's?" But this is not the content of the extra bad feeling that we had imagined as both natural and proper. Second, the propriety of the driver's regret does not depend on the harm he caused having been unintended, on its being a failure of one of his projects or an event in conflict with one of his values. It would be proper, albeit unlikely, for the successful murderer to feel worse than the unsuccessful one. (Nor is this thought wildly unlikely. For a person may pull the trigger and then think "What have I done?" In such a case, he may blame himself not for failing in a project, but for having had the project and succeeded.)

Unacceptable as this proposed answer is, two aspects of the analogy between the unlucky faulty moral agent and the unlucky spouse or scientist or ballplayer

are worth noting. First, these cases call to mind the variety of negative attitudes we may have toward ourselves that are not species of guilt nor primarily expressions of blame. We may be angry, frustrated, disappointed with ourselves in ways and for reasons that have nothing to do with moral fault or vice or harm. Second, these cases remind us of how natural and apparently appropriate it is for our attitudes towards ourselves to be affected by contingencies beyond our control. It seems natural and appropriate to be proud when one is awarded an honor or has proved a theorem, or has fixed a faucet without having to call in the plumber. It seems natural and appropriate to be disappointed or annoyed with oneself when one doesn't make the team or the shortlist, or when one screws up the faucet even worse. To be sure, one shouldn't be too proud or too self-critical on the basis of these successes and failures - the element of luck that enters into these things is considerable. But as we find fault with the person who takes too much credit or discredit for the successes and failures that are partly due to luck, we also find fault with the person who distances herself too much from all these events.

With these remarks in the background, let us return to the case of the unlucky driver, who has run over a child. This time, however, imagine that, moved by considerations that support the rationalist position, he is able to detach his feelings and judgments about himself from the unfortuitous consequences of his recklessness. "To be sure, I was at fault," he thinks. "I should have had those brakes checked last week. But thousands of others are similarly negligent every day. My behavior was no worse than theirs, and I am no more a 'murderer' than any of them."

Earlier, I suggested that such a response would strike us as appalling. I now want to consider why it would strike us that way. Perhaps we suspect that the driver is trying, or succeeding, in judging himself less harshly than he deserves - that he is disclaiming responsibility for an act or a harm for which he is objectively responsible. But I believe that there is something else that is disturbing, not about the level of self-assessment that his detachment from the accident makes possible for him, but rather about the detachment itself. There is something disturbing about the agent's thought that the child's death, sorry as he may be about it, has relatively little to do with him.

The concern is not that his thoughts or judgments are, strictly speaking, false. Indeed, they may not be. He is right that his negligence is no different from the negligence of thousands of others, and if he feels considerably guilty, and blames himself to a significant extent, then it may well be that his sense of his own blameworthiness is within the range that accuracy requires. What is problematic is his failure, beyond this, to take the consequences of his faultiness to have consequences for him, to be a significant part of his personal history, in a way in which witnessing, much less reading about an accident would not be. The problem is not that he refuses to accept what responsibility he objectively has for the child's death; it is that he fails to take responsibility for it, in a way that goes beyond that. He reveals a sense of himself - his real self, one might say - as one who is, at least in principle, distinct from his effects

on the world, whose real quality and value, for better and for worse, is at best impurely indicated but not at all constituted by the goods and the harms, the successes and the failures that comprise his life in the physical world. It is as if he draws a circle around himself, coincident with the sphere of his will.

If one contemplates this attitude, not just as a single reaction to a single incident but as a way, if you will, of being-in-the-world, one might well question the intelligibility of the conception of identity that lies at the heart of it. Here, however, I am concerned to make a more purely normative point - namely, that even if this attitude towards life can be made out to be conceptually coherent, it defines an approach to life that is unhealthy and undesirable.

A nameless virtue

Let me elaborate on the slightly paradoxical claim I made two paragraphs ago, that what is troubling about the fully rationalist truck driver is not that he refuses to accept what responsibility he objectively has for the child's death - it is that he fails to take responsibility for it, in a way that goes beyond that. It is this thought, and the virtue to which it refers, that I believe properly lies at the heart of the solution to the problem of moral luck.

There is a virtue that I suspect we all dimly recognize and commend that may be expressed as the virtue of taking responsibility for one's actions and their consequences. It is, regrettably, a virtue with no name, and I am at a loss to suggest a name that would be helpful. It involves living with an expectation and a willingness to be held accountable for what one does, understanding the scope of "what one does," particularly when costs are involved, in an expansive rather than a narrow way. It is the virtue that would lead one to offer to pay for the vase that one broke even if one's fault in the incident was uncertain; the virtue that would lead one to apologize, rather than get defensive, if one unwittingly offended someone or hurt him. Perhaps this virtue is a piece or an aspect of a larger one which involves taking responsibility not just for one's actions and their consequences, but for a larger range of circumstances that fall broadly within one's reach. One may offer to pay for the vase one's child broke, or offer to take the blame for the harm someone suffered as a result of the practices of an agency of which one is the head. Like other virtues, this one is a matter of offering the right amount (whether it be of compensation, apology, or guilt) at the right time to the right person in the right way. It is not the case that the more responsibility one takes for the harms that lie at increasing distance from one's control, the better.⁵ Yet one ought to take responsibility for more than what, from a bystander's point of view, would be justly impersonally assigned.

If I am right in thinking that this is a virtue, and one, moreover, that most of us at least implicitly recognize, then it gives us a way of understanding our responses in moral luck cases: Equally reckless drivers, equally negligent building inspectors, and so on, deserve equal blame for their faulty behavior. But those who cause harm are called upon to feel and to do something more that others are not similarly called upon to do and to feel. It is, in other words, the

occasion for the one agent to display his (nameless) virtue or lack of it, while it is no such occasion for the other.

The idea that we regard an expansive sense of responsibility (when costs are involved) as a virtue provides a coherent explanation of the pair of intuitive responses that earlier appeared to be in tension. Whether these responses are justified, however, depends on the further question of whether we ought to regard this as a virtue - whether, if you will, it is a virtue. There seem to me to be two reasons for thinking that it is.

Perhaps the more obvious reason for regarding it as a virtue is that, when applied to harmful actions, this trait is a species of, or at least akin to, the well-established virtue of generosity. Generosity generally involves a willingness to give more - more time, more money, more love, more lenience, more, in one way or another, of oneself than justice requires. In offering to pay for the broken vase, in trying to ease the pain or provide comfort to the grieving family beyond what a rationalist assignment of liability would demand, an agent voluntarily benefits or tries to benefit others at cost to herself. That this should be seen as virtuous is not hard to understand.

There is another aspect to the character trait in question, however, that is worthy of support as well. This aspect comes more sharply into view if we contrast a mere witness to the breaking of the vase who offers to replace it with the person who, innocently or not, actually broke the vase. In either case, the offer to pay would seem to be an act of generosity. But while we might appreciate and even admire the bystander's offer, we might also be slightly puzzled by it, or even, in certain cases, resent it. On the other hand, we are apt positively to expect the offer from the agent who broke the vase, and to be puzzled or disturbed by his failure to offer to pay. We expect the vase-breaker, like the truck driver who hit the child, to acknowledge that the consequences of his behavior have something specifically to do with him. We expect the vasebreaker to offer to pay, then, not only because we want him to be generous, but because we expect him to accept contingency in the determination and assessment of who he is.

While generosity is a thoroughly moral notion, this other aspect of the virtue I am discussing is not so clearly or narrowly moral. The reason for objecting quite generally to an attitude of greater detachment and for commending an embrace of at least some of what lies beyond the sphere of one's will has less to do with a benevolent concern for others than with a view about what, for lack of a better word, might be called psychic health. The desirability of this trait comes partly from its expressing our recognition that we are beings who are thoroughly in-the-world, in interaction with others whose movements and thoughts we cannot fully control, and whom we affect and are affected by accidentally as well as intentionally, involuntarily, unwittingly, inescapably, as well as voluntarily and deliberately. To form one's attitudes and judgments of oneself and others solely on the basis of their wills and intentions, to draw sharp lines between what one is responsible for and what is up to the rest of the world, to try in this way, to extricate oneself and others from the messiness, and the irrational contingencies of the world, would be to remove oneself from

the only ground on which it is possible for beings like ourselves to meet. If we define ourselves in ways that aim to minimize the significance of contingency and luck, we do so at the cost of living less fully in the world, or at least at the cost of engaging less fully with the others who share that world.

The moral of moral luck

In my introductory remarks, I mentioned that something seemed to be missing in the rationalist response to the phenomena that the concept of moral luck concerns, something that suggests a grain of truth in the irrationalist response. The conclusion to which the train of thought in this paper has been leading is that what is missing is an acknowledgment of the nameless virtue that urges us, as a matter both of moral character and of psychic health⁶, to recognize and accept (to an appropriate degree) the effects of our actions as significant for who we are and for what we should do. By explicitly including a description and endorsement of this virtue in our response to the phenomena at issue, we reach a position that we may think of as falling between the starker alternatives of rationalism and irrationalism as I initially presented them. With the rationalists, this position holds that equal fault deserves equal blame, and that, therefore, an impartial observer ought to judge equally faulty actions to be equally blameworthy. Yet, with the irrationalists, it allows that different effects call for different responses, so that the faulty - or, for that matter, the nonfaulty - agent's attitude toward herself as well as her deliberations about what to do from here are properly affected by her actions' effects.

This proposal is abstract and leaves many questions open. The proposal of the nameless virtue to which I alluded, for example, calls out for further refinement: How much and what kinds of difference in agents' emotional and practical responses should the contingent and uncontrollable differences in the effects of their voluntary behavior make? How, if at all, does this alleged virtue apply or appear in connection with the positive or beneficial effects of actions that are at least partly not traceable to the quality of the agents' wills? Should we restrict the scope of this virtue, or of the evaluative judgments relating to it, to the realm of moral luck in "how things turn out," or does the position sketched in this essay have implications for our responses in other areas of moral experience where we must also confront the considerable role played by luck (for example, luck in one's circumstances, or luck in one's natural moral endowments)?

These questions are intrinsically interesting, practically important, and apt, I suspect, to arouse much controversy. Without answering them, it may be said, we cannot claim to have "solved" the problem of moral luck or laid it, once and for all, to rest. Still, I believe that if we can agree on the basic features of the admittedly abstract position for which this essay has been arguing, considerable progress will have been made. We will be able to identify much more in common in the intuitions and theoretical positions that formerly may have seemed diametrically opposed. Disagreements will have been narrowed and conceptually sharpened.

The framework of the abstract position advocated in this essay leaves room, then, for many variations in response to more detailed questions concerning the range of phenomena involving moral luck. I shall end by discussing one pair of such variations that seems to me particularly curious, and that may be especially closely tied to a contrast in sensibilities that might formerly have attracted people to rationalist and irrationalist positions on moral luck respectively.

Blame, guilt, and agent-regret

The position for which I have argued states (in agreement with the rationalist) that blameworthiness is solely a function of faultiness. In other words, equal fault deserves equal blame. At the same time, my position holds (in agreement with the irrationalist) that different effects call for different responses - including different emotional responses in the agents whose behaviors bring about these effects. A question that arises is how the difference in appropriate emotional response is related to the equality in appropriate blame - for blame, it would seem, and especially blame of oneself, is itself an emotional matter.

One response to the question involves the claim that there are two distinct emotional responses at issue that sit, as it were, side-by-side. The emotional response to beliefs about one's blameworthiness is traditionally called guilt. The emotional response to beliefs about the badness of the effects of one's actions is what Bernard Williams labeled 'agent-regret.' According to this first position, guilt is the emotion one feels or should feel in proportion to how much one judges oneself blameworthy. Agent-regret, by contrast, is a special form of sadness or pain accompanying the wish that things had been otherwise with regard to something with which one's agency was somehow involved. This position allows us to be pure and simple rationalists with respect to the issue of blame, while also supporting the view that it is reasonable and appropriate that two agents whose acts turned out very differently should, in some way, feel differently about themselves and their acts. How much guilt one would feel, on this view, would be a function of how blameworthy one judged oneself to be; how much "sadness" would be a reflection of the nameless virtue.

This proposal neatly addresses the tension between the equality in appropriate blameworthiness and the inequality in appropriate negative feelings that depend on the consequences of one's actions by assigning the appropriate responses to separate emotional categories. To some, however, including myself, this proposal seems too neat. Our emotions do not seem to sort themselves out as clearly as this position suggests. This thought gives rise to a second proposal.

This position, like the earlier one, acknowledges a distinction between guilt and agent-regret. Guilt is at least partly a function of judging oneself to be blameworthy. In cases in which we judge ourselves to have done something wrong but which - luckily - have no bad effects, guilt, and not agent-regret, is an appropriate response. Moreover, guilt is not appropriate when we have done nothing wrong. In cases in which we are not at fault but which nonetheless - unluckily - lead to serious harm, agent-regret, and not guilt, may be appro-

appropriate. However, when a faulty action causes a harm, our feelings tend to get all mixed together - several sources for feeling bad about oneself and for being motivated to engage in certain sorts of actions combine to produce a mental state that is not itself analyzable into component states. Moreover, in light of a point I made earlier about the indeterminateness of the proper "amount" of blame to be assigned to a given act, this phenomenon seems logically as well as psychologically inevitable.

Earlier, I pointed out that although one can sensibly make some comparative judgments among wrongful actions, noting that one act is more blameworthy than another, and that some acts are very blameworthy and others only slightly blameworthy, it is unclear how fine-grained and precise such judgments can meaningfully be. Equally important, there is no apparent principle correlating degrees of blameworthiness with uniquely appropriate amounts of pain, punishment, or guilt. Because of this, it is inevitable that how one feels in connection with an acknowledgment of blameworthiness will depend on more than the degree of blame one acknowledges oneself to deserve. How much one blames oneself, in other words, must be a function of more than how blameworthy one thinks one is - for how blameworthy one is cannot supply a complete answer of how much blame one should get. Thus, the fact that the truck driver who kills a child blames himself more than the other driver need not indicate that he judges himself to be more blameworthy than the other, for how much he blames himself is a function not only of his judgment of blameworthiness but also of his reasons for agent-regret.

The recognition that the amount of blame and guilt that it is appropriate to direct toward an agent is significantly indeterminate - the recognition, that is, that blame and guilt cannot exclusively be a function of blameworthiness - may be combined with an endorsement of the nameless virtue to support this second response to the tension with which we have been concerned. Whereas the first response emphasizes the difference between guilt and agent-regret, this one emphasizes the complexity of the phenomena of blame and guilt - the different reasons, and reasonable factors that shape whether and how much blame we allot, both to others and especially to ourselves. This position would endorse the idea that one person might properly blame herself more than another whose action was equally faulty, and so it may be said to concede more to the initial irrationalist position than the previous one. At the same time, it retains the notion that lies at the core of the rationalist view that how much blame one deserves, or, perhaps better, how deserving of blame a person is, is purely a function of faultiness in action, and not at all a function beyond that of how one's actions turn out.

According to this view, then, moral luck - at least the sort of moral luck we have been discussing, luck in how our actions turn out - is a reality, rather than an illusion, in this sense: that how good or bad a person should feel about herself, how much she should blame herself (or pat herself on the back), how much or how little it is morally incumbent on her to do is in part a function of how her actions turn out in ways that are beyond her control. And yet it

is not, as the irrationalist position would suggest, because one's moral record, as it were, that establishes one's goodness or badness as a person is determined by these nonvoluntary matters. Rather it is because, as creatures of a physical and social world, we have among our responsibilities or perhaps less moralistically among our given projects, one that urges us to recognize the effects of our actions, as things that connect significantly to us, that have repercussions for who we are and what we should do.

The moral of moral luck, again

What is the moral of moral luck? I regret that I do not have a slogan of twenty-five words or less that sums it up beautifully. However, if I am right about my analysis and assessment of the phenomenon in question, then the moral is one that must recognize the considerable truth in both the rationalist and the irrationalist positions. It is that a morally conscientious approach to life, as well as a humanly conscientious approach to morality, must strike a balance between an interest, on the one hand, in attaining the kind of justice that comes from limiting the significance of that which is independent of the power and the quality of our wills with an interest, on the other hand, in acknowledging our earthly character and maintaining our connection to the social and physical world. The paradoxical quality of the phenomenon of moral luck comes from the fact that this latter interest can only be served if we affirm the significance that the former interest seeks to limit.

Johns Hopkins University

¹ Bernard Williams, "Moral Luck" in Bernard Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981) pp. 20-39. See also, Thomas Nagel, "Moral Luck" in Thomas Nagel, *Mortal Questions* (Cambridge: Cambridge University Press, 1979) pp. 24 -38.

² It is, for example, a function not only of how serious a moral breach is involved, which itself is a complex matter, but also of the degree to which the agent could have been expected to appreciate the seriousness of that breach, and how free she was to govern her actions according to her appreciation of its moral status.

³ There is another reason to be dissatisfied with this "linguistic" defense of the irrationalist position I have presented - namely, that the linguistic point cannot easily be made about other cases which seem morally analogous to the case of the reckless drivers. In the drivers' case, our language readily and naturally provides different descriptions of what blameworthy things the two drivers have done. But other examples with the same moral structure may not lend themselves to similar differences in description. Consider a woman who has a clandestine affair with a married man, whose spouse, upon discovering the deception, commits suicide, and compare her with the woman, who, having no more and no less reason to worry about her lover's spouse, sees the relationship through without the wife's ever knowing. As with the truck drivers, the women in these two cases might naturally and appropriately feel very differently about themselves. The one, after all, has been involved in some-

thing that has led to an innocent woman's death, and the other has not. But one wouldn't say that the lover in the first case had killed the wife, or even caused her death. We might say that her actions contributed to events that led to the suicide, but that description does not in itself make clear that she acted in a way that was blameworthy at all. (What triggers a suicide in a sufficiently troubled person may be totally innocent, after all.) The most natural description of what the first woman did wrong applies just as well to the second - she failed to take the wife's interests sufficiently into account. And yet, I submit, it is perfectly natural that, given the way things turned out, the one woman should blame herself more than the other.

- ⁴ Bernard Williams makes a similar point in the article that introduced this example, *op. cit.* (his point was addressed to the lorry driver whose driving was entirely faultless). The point that when we take the observer's point of view we are less sympathetic to the irrationalist position seems even clearer when we consider intentionally immoral acts and their equally faulty attempts. For example, in the case of the murderer and the attempted murderer, the inclination to judge and to blame them equally seems to me even stronger or more clearly strong, than in the case of the reckless drivers. Perhaps this is because the indeterminacy of blameworthiness (and also of badness and wrongness) is less problematic in the case of murder and attempted murder, than in the case of - pardon the expression - middle-of-the-road recklessness. We all know that murder and attempted murder are very bad indeed, and that a great deal of blame in either case is appropriate.
- ⁵ Thus, for example, it is far from obvious that it would be a virtue for all Germans to feel responsible for the consequences of the Third Reich.
- ⁶ From an Aristotelian point of view, there is no distinction between these values.