



CHAPTER 6

How Justice Pays

In this chapter and the one to follow I want once again to make connections between (humanistic) psychology and philosophy. But I will move away from questions of epistemology and concentrate exclusively on what might be called the philosophical intersection between moral psychology and ethics. The issues I want to deal with are ancient ones, though they continue to exercise ethicists and moral psychologists. They concern in some broad sense the justification of morality, of moral virtue or acting justly, and they fall roughly into two categories: questions about whether justice or virtue pays in egoistic terms and questions about the kind of challenge psychological egoism may be thought to present to moral philosophy/ethics. I shall make some previously unexplored connections between psychology and ethics in both chapters, but I want to begin in the present chapter with questions about the justification of justice. In our next chapter I shall bring out issues concerning the nature and moral implications of psychological egoism that lie just below the surface of what we are going to be speaking about in the present chapter. Our main question right now is whether justice pays, with the issue of how justice pays then emerging from the discussion of this first question. However, in order to deal properly with either question, I need to give you a fair bit of background.

1

The question whether justice pays originated, as far as we know in ancient Greek philosophy. (The Greek term for justice “dikaosune” can also be translated as righteousness or as moral virtue.) It originated against a certain background, the background assumption or view that we today call eudaimonism. On one modern understanding eudaimonism is the view that virtue in order to be genuine (or genuinely valid) virtue has to pay, in self-interested terms, for the virtuous person. It was then also assumed or argued that vice or injustice doesn’t pay in such terms. Eudaimonism dominated the thinking of ancient Greek ethics. To be sure, there were Greek ethical skeptics and even Greek nihilists about ethics, but among those who promulgated positive ethical doctrines or theories, eudaimonism was always assumed. Plato, Aristotle, the Stoics, and the Epicureans all assumed eudaimonism in one form or another. Thus the most famous of ancient (Western) books of philosophy, Plato’s *Republic*, centers around the assumption of eudaimonism. It attempts to show that justice or moral righteousness pays in self-interested terms because it assumes that that conclusion is necessary to any justification of justice as a virtue.

To be sure, there is a moment in the *Republic* where Plato seems to deny eudaimonism. He says that the philosopher will and should be willing to abandon his or her preferred life of philosophizing in order to help govern the State. Individual preference and happiness have to yield, in other words, to the larger interests of the State. This sound like the idea of self-sacrifice, a notion otherwise absent from ancient Greek and Roman philosophy, and for a moment, then, at least, a Greek philosopher seems to be assuming that it can be a valid demand of virtue and of justice that one give up one’s own greater good—which would mean that eudaimonism doesn’t hold as a general thesis. But this interpretation of Plato is controversial, and the main point is or should be that Greek philosophy in general and Plato in particular seem (otherwise) committed to eudaimonism. In the *Republic* Plato seems committed to showing two young Greek men, Glaucon and Adeimantus, that (contrary to what Thrasymachus says in the dialogue) justice is no sham because it can be shown to pay, in self-interested terms, for those who are just.

Today in the West we think we know better. Our modern philosophical ethics, whether in Kantian, in Intuitionist, or in Utilitarian form, assumes that the question what our duties are can be asked independently of asking or considering whether acting in accordance with duty is in our own best

interests. Indeed philosophers like H. A. Prichard in “Does Moral Philosophy Rest on a Mistake?” (*Mind*, 1912) want to insist very explicitly that it is a mistake to think in the manner of eudaimonism that we have to show that morality pays in self-interested terms if we want to show that morality is justified and valid for us. Similarly, Kant makes it quite clear that our moral duties can be justified independently of our felt need to assume (in what he calls “practical postulates”) that there is an afterlife in which God rewards the virtuous and punishes the morally vicious. Most philosophers today would agree with this dissociation of virtue and personal profit, but it took the West a long time to recognize that valid morality needn’t pay in self-interested terms.

However, Chinese thinkers have implicitly known this from the start. Confucius, Mencius, and the other great Chinese ethicists never assumed eudaimonism. Indeed, there seems to be relatively little effort in China to show that moral virtue can help one to have a better or happier life (as Li Jialian has pointed out to me, the *Yizhuan* is something of an exception); and the assumption that we need to show that virtue pays for the individual in order to assure the validity of virtue was never accepted much less considered in ancient China. Nor, as far as I know, was it accepted or considered during the much later period of neo-Confucianism. Somehow, the Chinese implicitly saw what was far from obvious to the ancient Greek and Roman philosophers: that a valid morality doesn’t have to be tightly tied to the self-interest or happiness of moral agents. This is something it took the West more than 2000 more years to recognize. Today we think that a really and seriously moral person will not and should not question their commitment or devotion to morality if an issue like eudaimonism is raised with them. This is Prichard’s explicit point in the article mentioned above (Kant is much less explicit about this issue), but then it would seem that such moral seriousness, such moral commitment/devotion, was present in China long before it emerged in the modern West. Chinese philosophy was admirably there long before Western philosophy got there.

2

But now the reader may ask why I am going to attempt to show here that morality in some sense pays if we already agree that we don’t need to show such a thing in order to validate morality or our own moral seriousness. We may have learned this more slowly in the West than it was at least implicitly understood in China, but now, at least, we are all or almost all

(there are a few ethical egoists around even today) on the same page about the necessity or importance of vindicating morality in self-interested, that is, completely egoistic terms. We understand that morality grants a limited but large space to self-interest, but we all or almost all are now clear that morality can validly go well beyond self-interested motivation or results. So is there any reason for us to continue on with the ancient Greek project at this point?

This very question, however, betrays a misunderstanding of what is at issue, ethically at issue, here. From the presumed fact that morality doesn't require vindication in egoistic terms it simply doesn't follow that there is no philosophical or human interest in determining whether morality pays in such terms. My teacher Philippa Foot accepted a form of eudaimonism in her early article "Moral Beliefs," but she subsequently dropped that assumption while at the same time, and for many years, relentlessly pursuing the question whether virtue or justice pays. I say relentlessly because for some decades she produced different versions of an article—or were they different articles?—linking virtue with the happiness of the virtuous, and then after all those intellectual struggles, the book that finally emerged, *Natural Goodness*, essentially gave up on the original project, offering no clear-cut or univocal vindication of morality in self-interested terms.¹ In any event, though, Foot never seems to have given up on *wanting* to connect virtue and happiness or a good life for the virtuous individual. Even while conceding an ambiguity or unclarity regarding the arguments that can be given for and against the tie-in between virtue and happiness, her language in *Natural Goodness* shows a great deal of frustration at her inability to say more by way of making the long-sought (long sought by her and of course much longer sought by Western ethicists) connection between virtue and happiness.

But now the reader will want to ask *why* she should have been so interested in making that connection and why I too am interested in doing so, given the rejection of eudaimonism. May it not just be a student following a teacher over a cliff? Well, I cannot speak for Foot and she is no longer with us to be consulted on this issue, but I can speak for myself and relate my motivations in this area to considerations that more generally support the idea of linking virtue and happiness in what we could call a non-eudaimonistic world. A lot of philosophers are interested, still interested,

¹ See "Moral Beliefs" in Foot's *Virtues and Vices*, Oxford: Blackwell, 1978; and her *Natural Goodness*, Oxford: Clarendon Press, 2001.

in understanding whether and/or how virtue and self-interest can be related, and I think I can explain to you why such an interest, and my own exemplification of that interest, doesn't have to be connected with any acceptance of eudaimonism.

Think about it! Why shouldn't a philosopher who understands how moral seriousness involves a lack of commitment to eudaimonism and who lacks such a commitment herself or himself not simply be interested, philosophically curious, about whether morality pays in self-interested terms. Their moral seriousness may well mean that they won't give up on morality (whatever that may mean—see Chap. 7 below) if it turns out that it requires a good deal of self-sacrifice, but people and philosophers as people do have an interest, at least somewhat independent of morality, in having a good or happy life, so won't those who are serious about morality be interested in the conditions of their own happiness? And won't at least the philosophers among them be interested in finding out whether these two things they are interested in and take seriously—their own happiness and morality—are closely related? Won't they want to know, most particularly, whether both these objects of their profound interest can or are likely to occur together in their lives? The answers are yes, yes, and yes, to the three just-concatenated questions. In effect, then, I am arguing that any morally serious philosopher with a modicum of self-concern could raise and be interested in answering the question whether morality/virtue/justice pays in self-interested terms, that is, whether it is likely to make one happier or better off in life than being immoral would.

But now we need to be frank about another matter. We are interested in our own welfare, but that doesn't prevent us from being interested in the welfare of others. (More on this in Chap. 7.) However, we do tend to resent those “anti-social” individuals who seem to lack any concern for others and who inflict a good deal of pain and suffering on other people. If it turns out that such immoral individuals are likely to have less good lives than those of us who are serious about morality and therefore concerned about the welfare of other human beings (and about keeping their promises, etc.), that fact is likely to please us. We think it would be unfair for them to benefit from their wrongdoings or overall have good lives, and would be pleased to learn that this isn't likely to happen. All the more reason, then, for us or those of us who are philosophers to try to find out whether what we would like to happen in this respect is in fact likely to occur. The desire to see certain people not benefit from their moral crimes against others is perfectly consistent both with normal self-interest (though

as Chap. 7 will make clearer that desire is itself not a self-interested one) and with ordinary moral willingness to make sacrifices. If I can show that highly immoral people don't have good lives, that will be a fact of interest and with pleasing implications for ordinary people, for philosophers, and, of course, for me. There is all the more reason, then, to attempt to demonstrate, as I shall attempt here, that psychology gives us reason to believe that the morally worst people (I will explain this specific idea in what follows) don't have good or happy lives. Our interest in these issues doesn't at all depend on assuming eudaimonism.

The idea that an appeal to psychology can help us show that justice/virtue pays is, of course, not a new one. Plato made various important appeals to psychology in arguing as he did in the *Republic* that justice pays. By justice or righteousness Plato meant that the parts of the soul or psyche were kept in order or ruled over by the part of the soul he called reason; and he got this conception of individual virtue or justice from an analogy with justice in the state. That analogy is strained, to say the very least, but using his notion of individual justice Plato argues (very roughly) that an unjust soul not ruled by reason will have insatiable appetites and end up frustrated and unhappy. This is certainly an appeal to psychology and it has a certain plausibility. If one is extremely immoderate in one's appetites, if one's appetites know no limits, one is unlikely to be very satisfied with or in one's life.

However, it is one thing to show that an immoderate insatiable person will not be very happy, quite another to show that someone who is unjust in the most ordinary ways, someone who steals, rapes, cheats, and murders is invariably immoderate and incapable of keeping their appetites within rational bounds. Hume's "sensible knave" (as per his *Enquiry Concerning the Principles of Morals*) might, for anything Plato shows, be both unjust and moderate in their appetites/desires, and Plato offers us no argument to show that such a person is impossible or incapable of happiness. So there is reason to be unsatisfied (sic) with Plato's attempt to show in psychological terms that justice pays. Still, we might look elsewhere in psychology to make the connection Plato sought to make. I propose to do just that, but the history of philosophy (and it is only philosophers who attempt to show that justice pays [in this life]) doesn't give us much to base any new psychological account on.

Some of those who have argued that justice pays (Foot among them) have spoken of the difficulty of escaping detection if one acts in an anti-social fashion and of the consequent ostracism or punishment an immorally

acting individual is likely to face. This familiar line is closer to sociology than to psychology, and it rings true *to a certain extent*. Even though some criminals and cruel people have escaped detection or have risen to such power that detection doesn't undermine their position or the good things they enjoy, still the chance of getting away with such lifelong immorality is relatively small, and that is one reason to think that the worst kinds of immorality don't pay. That still doesn't tell us whether the virtuous individual is likely to do well in life or better than the immoralist, but it is something worth knowing and telling. Still, I would like to be able to say more, and in order to do that I am going to have to go in a psychological direction that is novel in relation to the issue of whether justice pays. This new direction requires, further, that I limit my aspirations to showing that the worse kind of immorality tends to rob such individuals of happiness or contentment. I won't have any separate and positive argument to show that the virtuous are likely to do better in life than the morally worst individuals, but *if we can show that the worst individuals are likely to be unhappy, then in the total absence of any parallel argument in relation to the virtuous or just, there will be reason to think that the latter are likely to do better in life than the most immoral individuals are*. That, at any rate, is as close as I will get to arguing or showing that justice pays, and it is good in any event to have a reason, additional to those previously offered by philosophers, for thinking that the morally worst individuals are unlikely to be happy in or with their lives.

3

So what are the new psychological considerations I want to bring to bear on the (roughly conceived) issue of whether justice pays? And do these considerations help us to see *how* justice pays, or, more accurately, *how* total immorality is likely to deprive someone of real happiness? Before we seek to answer those questions, however, we need to consider with some care what it is for someone to be totally immoral, to be among the morally worst of human beings, and to that end it will help if we can come up with some sort of typology, not of general human types, but of moral types. Here the moral typologist par excellence can be of help to us. James Martineau's *Types of Ethical Theory* (Oxford, 1895) not only gives us a typology of ethical theories but also offers its own take on the different moral types of individuals. It does so somewhat obliquely, but for our purposes here, what Martineau offers is a commonsensically, intuitively,

plausible take on what kind of person is morally best and what kind of person is morally worst.

According to Martineau, the best kind of person (understood in secular terms) is someone who is compassionate and benevolent or, as we say in today's vernacular, caring. The worst kind of motivation, according to Martineau, is malice, and it is not much of a jump from that idea to the conclusion that the malicious person is the morally worst kind of individual. I don't know of anywhere else in the literature of ethics where these "moral endpoints" are so definitively stated and marked. However, with these endpoints as given, one can wonder whether there is any such thing as a moral mid-point, and Martineau gives us absolutely no help with that issue. And it is an interesting issue because it turns out to be interestingly ambiguous.

One would be inclined to think of the midpoint in a plausible moral typology as something or someone who is morally neutral, but the notion of neutrality here can be given two very different interpretations or explanations. Malice (or maliciousness or malignity) aims at harming other people and benevolence, compassion, caring (concern for others) aim at helping them, so neutrality might then naturally seem to be a state of mind or motivation that has neither of these aims. We have a word for such an attitude, for such motivation: we call it *indifference* to others (other people or animals), and in purely conceptual terms it makes a certain sense to see indifference as poised or placed halfway between malevolence and benevolence. However, if we do this we must recognize that we have *strayed away* from neutrality in a very definite different sense because indifference to others is far from being morally neutral. What are we to make of all this?

Well, I am inclined to say that indifference is indeed at the halfway point between malice and benevolence, but that that halfway point is not one of moral neutrality. Being neutral as between malice and benevolence is a definite state of mind and that state of mind is not a morally neutral one. To put it clearly but paradoxically, being neutral vis-à-vis the welfare or ill-fare of others is not a morally neutral attitude. Rather, it is an attitude we can morally criticize, even condemn, but we also of course and consistently with this point have to recognize that, commonsensically speaking, neutrality in the form of indifference to others is less morally bad than malice and morally worse than attitudes/motives of benevolence and compassion. So there is an ordering of morally better and worse among the three motives of benevolence, indifference to others, and malice, but

this ordering is asymmetric because what is arguably the midpoint among these motives is morally bad, so that we end up with two basic bad categories and one good one. In this abstract sense or way, it is harder to count as morally good than to count as morally bad.

I don't believe these points have appeared elsewhere in the literature of ethics, but I should immediately add that they are far from giving us a full moral typology. For one thing, what we have said doesn't tell us about the moral status of self-interestedness or of character traits that relate to the interaction between compassion and benevolence, on the one hand, and justice and deontological restrictions (on pursuing the good), on the other. We need to talk about these further issues now, and let's start with self-interestedness.

Self-interestedness is compatible with benevolence or compassionate concern for certain others and so is not the same thing as indifference to others. Nor, and for the same reason, does it involve being exclusively concerned with one's own self-interest/welfare. We shall see in the next chapter that total self-interest, that is, psychological egoism in the traditional sense, is not possible for us humans or even, probably, for any possible intelligent being. But leaving that point aside, it is important to realize that self-interestedness is not the same thing as *total and exclusive* self-interestedness, and, because self-interestedness simpliciter leaves much room for concern about the welfare of others and by no means entails malice toward others, self-interestedness really is morally neutral in a way that indifference to others is not. To that extent we can say that self-interest is a morally better motive than is indifference to others. (Obviously too, self-interestedness is not the same thing as selfishness.) What about justice and deontology? How much difference do they make to our typology as expounded in terms of four different moral types so far?

Let's start with what I think is the easiest case or pair of cases, the case of someone who kills one person to save three and the contrasting case of someone who refuses to do that, and let everything else be equal. Here deontology and consequentialism definitely disagree about what it is morally best or obligatory to do, and let's start with the perspective of consequentialism. The consequentialist will see the person who kills one to save three as falling within the basic category of caringness or benevolence and will thus regard such a person as potentially of the highest moral type. They will likely, on the other hand, criticize the person who refuses to kill in such circumstances as a rule worshipper (I am talking about act-consequentialism here). By contrast, the deontologist will criticize the

person who kills and treat the person who refuses to as potentially of the highest moral type. So it would seem that deontologists and consequentialists will advocate different moral typologies where this is conceived as involving a hierarchy of morally better and worse moral types. But I think they can both agree that when issues of deontology are not in question, benevolence is the best motive, and I think they could further agree that malice is the worst of motives, for even the deontologist could agree (and if they are enlightened and morally sensitive I think they would agree) that malice is worse than breaking the deontological restriction on killing in order to save a greater number of lives. For all we have seen, then, the status of malice as the morally worst of motives and of the malicious person as the morally worst kind of person can be maintained without much reason for philosophical or ethical dissent, and that is the main or a main point we need to rely on in the arguments I propose to offer in what follows.

However, to make those arguments we also need to become clearer about what malice actually (or necessarily) is. I want to argue that malice involves certain psychological underpinnings and cannot exist simply and utterly on its own. If it could, any connection between being malicious and being unhappy or discontent could probably not be made, but if we consider the psychological underpinnings or bases of malice, we will be in a position to recognize the possibility, and I shall argue the actuality, of a strong connection between malice and the unhappiness (or general discontent or dissatisfaction) of the malicious, malevolent, or malign person.

4

It is tempting or has been tempting to think of malice as something that can exist all on its own. For one example, in his classic literary study *Shakespearean Tragedy* A.C. Bradley spoke of Iago's "motiveless malignity" toward Othello, but I have doubts about whether this really makes good sense and whether it even makes sense in relation to the character Iago in Shakespeare's *Othello*. The play itself makes very clear that Iago's malice/malignity toward Othello is at least partially explained by the fact that he had previously been passed over for promotion by Othello. His malice may be disproportionate to the wrong or injury he had thus received at Othello's hands, but it is one thing to conceive a given instance of malice as disproportionate to a received injury (and here, to anticipate our later discussion, factors like paranoia can sometimes enter the picture) and it is quite another to suppose that someone's malice is motiveless, that

is, that there is absolutely no imagined or believed in reason for someone's malice. It is the latter that I want to suggest is impossible, but to make the case for that conclusion I think we need to refer to and make use of another sort of example, one ultimately derivable from the work of G. E.M. (Elizabeth) Anscombe.

In her 1957 book *Intention*, Anscombe pointed out that certain putative desires don't make intelligible sense and could never exist. Her clever but deep example was the desire for a saucer of mud. One could want saucer-shaped mud in order to give oneself a facial, that is clear. But Anscombe argued that no one could want a saucer of mud for no further reason beyond that very desire. A basic desire for a saucer of mud makes no sense and cannot be clearly imagined to exist.

I hope the reader gets the point of what Anscombe was saying here, and it is my belief that the point transposes to malice. Revenge by its very definition doesn't stand alone as a desire but is based on something a person wants to take revenge for, but malice doesn't wear its conceptual incapacity for standing alone on its face. One has to do some thinking and arguing before one can see the connection (assuming there is one) between what Anscombe said about the desire for a saucer of mud and malice conceived as a desire to cause (certain) others pain or harm (death being the worst of harms, presumably). There are potential objections to the idea that malice cannot stand alone, objections for which there are parallel objections to the idea that there cannot be a basic desire for a saucer of mud, and we need to consider these.

Perhaps the most forceful or forceful-seeming objection comes from scientific thinking about the causes of behavior. We know that human beings have certain built-in behavioral tendencies: the knee jerk is a good example. But there is no further psychological reason why one's knee jerks when it is struck in a certain way by the doctor. The knee jerk is a kind of automatism, and one might wonder why there couldn't be a malice automatism built into certain people. Now this sort of possibility can be fairly quickly countered via the consideration that a knee-jerk reaction doesn't require anything we can naturally call motivation on the part of the person whose knee jerks. It is not as if (we think) there is any desire there, much less some sort of basic desire, to move one's leg when one's knee is struck in a certain way. In fact, we wouldn't call the reflex voluntary or therefore treat it as an automatic kind of *action* on the part of the individual whose knee is struck. Malice and even motiveless malice is supposed to represent or constitute a desire to do, voluntarily do, harm to people, so this is a far cry from a mere reflex and clearly cannot be understood in terms of the idea of a reflex.

But the rejection of the example of a reflex can move us toward other cases that would seem to offer more of a challenge to the idea that malice and the desire for a saucer of mud cannot be basic desires. What if (certain) people were built in such a way that they always wanted to harm others, always sought ways to do that, and always performed such actions when circumstances allowed. We might speak, not of a reflex, but of a basic instinct to harm (certain) others, an instinct, moreover and as we are conceiving it, that was independent of means-end rationality: one is not trying to hurt others in order to procure some good for oneself or loved ones, but simply, and without further reason(ing), has an instinct for harming the way we are supposed to have, say, an instinct for survival? Why isn't such a thing possible, and if it is, doesn't that show that there could be such a thing as motiveless malignity or a motiveless, that is, basic, desire for a saucer of mud or saucers of mud?

But would such a putative instinct really be malice, a basic and strong desire to hurt people. I don't think so. The example as described hasn't built enough into itself to make it count as a case of basic human malice? All we have with this example is an individual that always tries to hurt others and takes all available means to doing so, unless she/he is threatened with danger to himself or herself, and so on. This doesn't show that the person or individual has malice toward those others or really wants to harm them, because it doesn't tell us anything about the emotions surrounding the supposed desire to hurt others. Usually, when we strongly want to do something, we feel hopeful or fearful about whether we will be able to accomplish what we want to accomplish. Further, we will be disappointed and even unhappy if we fail to achieve our end, but glad or happy if we do achieve it. But none of those emotions has been built into the description of the supposed example showing that motiveless malignity is possible. (Similarly for any parallel argument and example relating to the supposed desire for a saucer of mud.)

Indeed, as far as the example has so far been constructed, no emotions are being thought of as attached to the supposed instinct for harming people. It is simply supposed that if the individual with the instinct fails to harm someone at one time, they will renew their efforts to do so or redirect them toward someone or someones else. The ordinary but entirely conceptual connection between desire and emotion is left out of the picture or is somehow assumed away, but we have no right to do that if we are to soberly and accurately consider what is conceptually possible here. The example of an instinct to hurt or harm others that is supposed to

show us malice as a basic instinct doesn't show us malice, is much too thin in its psychological accoutrements to allow for that. In fact, the example as constructed treats instinctual malice and malice itself therefore as a kind of automatism. Not to be sure a mere reflex kind of automatism, but an automatism nonetheless because the example is devoid of the emotional dispositions that characterize desire as desire: an automatism, in other words, because the example fails to give us an example of desire, of malice as a desire, and leaves us only with some kind of mechanism, psychological or otherwise, that ensures attempts to harm or behavior that harms others, without bringing in what is most characteristic of human action or behavior: its basis in desires that centrally involve emotional dispositions. People and even philosophers can *say* that there can be basic malice, malice based on nothing deeper or further than the malicious person thinks or feels, but I don't think it is actually possible to construct or conceive an example where such a thing is fully realized. Basic malice really seems to be impossible.

But then we have to ask what sorts of things, or reasons or whatever, malice does depend on, and at least part of the answer is already staring us in the face. Iago's malice toward Othello constitutes or is based in a desire for revenge against Othello for having passed him over for promotion, and indeed the clearest cases of malice do in obvious ways involve a desire for revenge. In that connection, though, it is interesting to consider what Bishop Butler in his *Fifteen Sermons Preached at the Rolls Chapel* says and also doesn't say about the connection between malice and (the desire for) revenge. Butler argues (brilliantly and to my mind convincingly) that both revenge and malice, involving as they often do cutting off one's own nose to spite one's face, are non-egoistic motives. This idea has been challenged, but I hope to show you in Chap. 7 that Butler was correct on this point.

However, Butler never says that malice has to be based in revenge. He takes no stand on that issue and, more generally, on the question whether malice can be self-standing or basic to an individual's psychology. But I want to say and in effect have argued that if we take our lessons from Anscombe, we have reason to think a basic desire to harm others is no more possible than a basic desire for saucers of mud. The questions we now have to consider, therefore, are whether malice has any other source or basis than a desire for revenge and, if doesn't, whether the desires for revenge experienced by the morally worst individuals are inevitably or usually tied to certain forms of unhappiness.

5

We are finally ready to bring in the psychology that, as I think and shall argue, is relevant to the question whether the worst people, the most malicious people, are not happy in their lives. I am not going to bring in the actual literature of recent psychology because I don't believe any of the literature directly addresses the main psychological issue that faces us here, the question of whether in terms of a realistic understanding of human psychology, one can argue that malicious people are unhappy. What I hope to show or at least give you reason to believe is that there are two basic and intuitively realistic models of what lies behind malice. Both involve revenge, but the revenge involved is differently related to human emotion in the two basic cases.

And I shall make another assumption in order to facilitate the argument and in order, too, to engage realistically with the immorality I am going to be talking about. There are different degrees or levels of malice and revenge, and the revenge my older brother takes against a schoolmate of mine who has bloodied my nose is not comparable in moral seriousness and in degree of malice with the more extreme versions of malice that characterize the people we think are most morally bad or vicious. If my brother merely bloodies the nose of the boy (or girl?) who has bloodied my nose, that is revenge and embodies (at least temporary) malice to the extent, as it is realistic to suppose, that my brother thinks the person who bloodied my nose *deserves* similar treatment. He wants that person to suffer not as a means to some further good for himself (my brother) or me, but because, as he thinks, the person deserves to suffer, and that counts as malice even if it is also somewhat excusable or understandable or, on certain views (an eye for an eye and a nose for a nose), justified. But all this, as we can say, is *minor malice*, and I think we can assume that the morally worst people engage in malicious actions of a much more serious kind and on a much larger scale. They hurt other people, they kill other people, they rape or torture other people, and take pleasure in doing so apart from and independently of any (other) benefits they derive or may hope to derive from doing so. We are talking, in other words, about (serial) killers, about (serial) rapists, and about serial torturers who take sadistic pleasure in torturing their victims, and actually all of these categories involve some kind of sadism commonsensically understood. This is even, though in a more minor and temporary way, true of my brother's attitude and actions toward the person who has bloodied my nose. Malice, revenge, and sadism

are arguably part of one package, but this assumes that malice not only cannot be free-standing but also always depends on some sort of desire for revenge.

In what follows, I shall describe and discuss further the kinds of individuals I think and I believe most of you think belong to the category of the morally worst type of human (or other) individual. I think these individuals fall into two classes based on the somewhat different etiology of their malice and malicious actions, but in both cases revenge will play a role, and I think we have no clear or realistic conception of any other kind of individual whose extreme moral viciousness is constituted by malice based in something independent of the desire for revenge. But, as I say, the two basic kinds of immoral revenge-malice have somewhat different etiologies and the difference in etiology is related to their emotional life. Some malice embodies a desire for revenge based in or on fear and some embodies a desire for revenge based in or on anger rather than fear.

When we think of the morally worst individuals we know of or read/hear about, they are people who kill or torture, and so on, mercilessly. In other words, we are talking, roughly and as I indicated earlier, of serial killers and rapists and the like. But before we proceed any further with describing such people and their underlying psychologies (and how these latter indicate or make for unhappiness), I think there is a moral issue we have to face in connection with the kind of malicious person I have been very roughly describing. I have said that such people are the morally worst people we know of, but aren't they in many or most instances also psychopaths, and can we consider psychopaths morally responsible for being the way they are? (Sociopathy is either the same thing as psychopathy or a milder version of psychopathy. In either case, we don't have to talk about sociopathy any further here.)

If psychopaths are not responsible or thus blameworthy for their malicious tendencies and actions, then there may be a problem with our assumption that the serial killers and rapists are the morally worst kind of people, since, arguably, many of them are psychopaths. But then again there may not be as much of a problem as might initially appear. First of all, it is possible that psychopaths can validly be held responsible for their terrible actions—many courts of law treat them just that way. Second, even if they are not responsible or blameworthy for being the way they are and doing what they do, we may still feel strongly that what they do is wrong, horribly wrong. It is wrong to serially rape young boys and the person who does this is a kind of moral monster, the worst kind or one of

the worst kinds of people we know of. That is a judgment that we most of us strongly want to make, and in that case we may want to say that such people are morally vicious in the worst way (we know of—there may be unknown unknowns here) *even if* we cannot say they are blameworthy for being as they are. As we shall see at greater length in what follows, some serial killers have suffered horrible kinds of childhood abuse; and that fact may make us hesitate or more than hesitate to blame them for being as they are, if we think the early abuse *turned them into moral monsters without their having much say in or control over the matter*. So perhaps such people are not morally blameworthy for what they are, but we may still want to say that what they do and are is morally terrible or bad in the worst way we know of. There is a long tradition of distinguishing moral and ethical judgments of wrongness and badness or viciousness from judgments of blameworthiness. But I don't propose to go into that literature here in part because there is an equally long literature taking issue with that way of seeing things and (conceptually) tying moral criticism and badness to blameworthiness in a strong way. But I don't think we need to enter into or cite either literature here because I think it is fairly clear what we need to say at this point independently of who is right on the issue just mentioned. Perhaps, despite the childhood abuse certain psychopaths are properly held responsible for their deeds and viewed as blameworthy for what they have done. Perhaps they are not blameworthy because of serious abuse they suffered as children. Either way, we can still maintain that what they do is wrong, horribly wrong. It can be wrong despite the fact that they can't be blamed for being the kind of person who does such horribly wrong things, or it can be wrong compatibly with seeing them as adults who as adults are responsible for what they intentionally and maliciously do as adults independently of how they were treated as children.

To that extent, the issue of psychopathy may be a red herring. Even if the most malicious individuals are psychopaths, they still can count and seem to most of us to count as the morally worst type of individual, and therefore our present project needn't consider or further consider the issue of the moral blameworthiness of psychopaths or, for that matter, the question of how many malicious anti-social individuals actually count as psychopaths. So let us move on.

6

The main question we now face concerns the psychological character of totally malicious individuals, the sort of individuals who kill, rape, torture and seemingly just like doing so. From what I think most of us know about our society and about the varieties of people there are in the world, I think we can see or come to see that the totally malicious come in two main categories. Roughly, they include, on the one hand, paranoid individuals who fear that the people around them are out to get them and, on the other, people who are taking a kind of revenge for what they suffered earlier in their lives. Both these types are motivated by malice, but in the former case that needs some showing.

Some might say that the paranoid individual who fears the people around them (irrespective of who those people are and what they are actually like) is acting egoistically when they act against those feared individuals. If I fear your intentions, then even if that fear is paranoid and unwarranted, what I do to prevent you from hurting me is an instance of self-interested motivation. By contrast, and as Butler told us three centuries ago, malice involves wanting to hurt or harm others even if that is not necessary to one's self-defense or to gaining other goods just for oneself. Malice thus (properly) understood seeks to harm someone independently of any ulterior or practical motive beyond the doing of harm—and revenge is very much like this. But what we have said about those who kill out of paranoid fear of others seems to fall under the category of egoistic self-interest and therefore not to involve malice as we understand the notion.

In fact, though, that appearance is misleading. Yes, there can be an egoistic aspect or dimension to the killing of those we think are out to kill us. We thereby really do seek to prevent them from killing us and thereby satisfy or attempt to satisfy the presumably egoistic desire to go on living. But when the paranoid individual kills or harms those they think are out to kill or harm them, the motivation involved is standardly far from being exclusively egoistic. Just think for a moment about ordinary human psychology. If I learn that someone has tried unsuccessfully to hurt me and is no longer even capable of doing me any harm, don't I nonetheless have some tendency, perhaps a strong tendency, to want to do them some harm, and isn't that desire for revenge, one's malice toward that person, independent of any harm one has actually suffered and connected, rather, to what the other individual has *sought* to do to us?

With paranoid individuals the case is relevantly similar. Yes, they want to eliminate a perceived or imagined threat to their lives or their welfare, but that is not the end of the story. The paranoid person who fears others around him and thinks they are plotting to hurt him and/or actually making efforts to do so will also want to revenge himself on those people for what they are trying or have tried to do. He will think they deserve to be punished for what they intend to do to him in all his self-imagined innocence and will want to be the one to mete out the punishment, thinking that it is entirely appropriate and fitting that he should be the one to make them suffer. So I think revenge is basic both to paranoia-based malice and malice based in anger and indignation at what others have made one suffer (as a child). In both cases there is anger and action based on anger, and the main difference between them is that in the paranoid cases fear is a major element, whereas in the other cases it seemingly is not. But now I want to describe these two kinds of malice/revenge in a bit more detail and use the detail and the basic points we have made so far to tie malice of both kinds to unhappiness. Let us start with malice based in the horrendous mistreatment of a child.

Not every child who is cruelly mistreated turns into a moral monster. Sometimes, though it has not been univocally clear to psychologists how this happens, a child who has suffered develops a strong empathic sensitivity to the sufferings of others, especially children. But that is not the usual case, not by a longshot, so let's talk about what more usually happens. If a male child has been repeatedly raped by a family member during their childhood, they can become a serial rapist themselves, as adults seeking to do to other children what was done to themselves. This is an instance of what psychologists sometimes call identification with the aggressor and what in popular speech is sometimes called Stockholm syndrome. The American public is perhaps most vividly aware of this syndrome in the instance of heiress Patty Hearst, who decades ago was abducted by the Symbionese Liberation Army and notoriously became a loyal gang member.

But I think revenge for childhood abuse more typically or famously takes the form of displaced aggression against those who abused the child. If the mother abused the child, then the child can become a serial rapist or murderer of women who roughly resemble the mother as she was when the abuse was meted out. (If the mother lets her brother or a stepfather rape her son, it may still be she and not the uncle or stepfather who becomes the symbolic target of later aggressions.) So much, then roughly, for etiology, but now we must consider how the child who identifies with

the aggressor or later seeks out symbolic targets that resemble their former abuser feels about what is happening or has happened to them.

Imagine yourself there, at the mercy of some adult, as a child needing and wanting love and affection, but receiving physical or sexual punishment/abuse instead. What a horrible situation and how horribly you are going to feel about it! Such children suffer and suffer greatly at the hands of those who are supposed to take care of them, and do you think such misery is easily forgotten or papered over once the abuse stops (later in childhood or in adolescence as the child becomes an adult)? It may not be pleasant to have to say it, but children are like other people, like adults. Although they may not have the adult vocabulary in which resentment can later on be morally expressed, they may still resent what is being done to them or has been done to them. And resent is too weak a word. They can be furious about what has happened to them, in an unappeasable rage at their suffering and mistreatment, and that unappeasability may come out later in the form of revenge, revenge against those who resemble the people who abused them or, twisted by identification with the aggressor, revenge against children who resemble their own abused selves. But the whole basis for this is the extreme misery of the child, and if the child becomes an adult who, finally, can take some sort of revenge, can the extreme misery have simply vanished from their mind? Isn't it more plausible to say that such misery has finally found an outlet, and if the misery needs, still needs, an outlet, then surely it has *not* vanished from the mind (or xin). How can such a person be deemed contented, happy, or pleased with their life?

So, you see, (a major portion of) the morally worst people we know of, the serial killers and rapists and the worst kinds of sadists, cannot be regarded as well-off in their lives. People who are not like that may die young or be subject to crippling diseases—their lives may also not be good ones. But those who are not afflicted in those ways and who are not full of vengefulness stemming from childhood abuse can have decent or good lives, and that constitutes an answer to our narrowly focused question as to whether justice or virtue pays. Now I haven't specifically been talking about the paranoid type of vengefulness, but what I have just been saying applies equally, I think, to them.

Abraham Maslow tells us that people want to be liked, esteemed, even loved. I shall say more about this in Chap. 7, but for the moment the point to take from his ideas is that someone who thinks everyone is plotting against them cannot be very happy about that fact. And then, as I indicated

earlier, such a person will be angry and vengeful toward those they think are plotting or have acted against them. But when you think of it, living in or with anger is not a pleasant way to be or to live. So again and on these further grounds we can say that the paranoid killer—someone, for example, like Stalin who (via “show trials”) kills off imagined enemies using their dictatorial powers—is not living a pleasant or contented or happy life. These people are further denizens of the category of morally worst people. But all the people of this category I am aware of lead less than happy lives. They are too angry, too unhappy with what is being or has been done to them, to lead a life we can think of as well worth having and living.² However, that doesn’t mean that the vengeful and malicious morally worst type is aware of how miserable or lacking in happiness and contentment they are. They may not be aware of how morally bad they are (paranoid individuals always feel justified), and by the same token, given our arguments, they may not be aware of how unfortunate their lives are.³

²Here I am assuming the obvious fact that the vengefulness of the worst people cannot fully slake itself through some single act of rape or murder. That is all the more reason to think that those people cannot, except perhaps temporarily, forget what has made and will continue to make them unhappy.

³I think it is worth noting the parallels between what I have been saying about very malicious very vengeful people and some more familiar and less unpleasant phenomena. If someone loses a valuable ring, looks everywhere for it, and eventually finds it, they will be relieved to find the ring and may idiomatically describe themselves as happy to have found it. But no one would wish to temporarily lose a ring in order to have the pleasure of finding it, and Plato says something similar about itches. It may be a relief or pleasant to scratch an itch, but no one would want to have an itch in order to experience the pleasure of scratching it. By analogy, one might grant that vengeful people feel a certain pleasure and perhaps glee at taking revenge and even at the thought of taking revenge, but a life of being mistreated and pleasurably taking revenge is no more good or pleasant *on the whole* than is the total experience of losing and finding or of itching and scratching. The ancient Greeks and Buddhists have gone on, however, to argue for a general pessimism based on the supposed unpleasantness of all desire. If desire is unpleasant, then even if one takes pleasure in satisfying it, one’s total experience is, according to this all-too-familiar view, unpleasant on the whole. Since human life is shot through with desire, the conclusion the Greeks and Buddhists drew (or, in the case of the latter, draw today) is that human life is on the whole unpleasant. But nowadays I think we are more savvy about desire and realize that not all desire is unpleasant. When one looks forward to a great meal, one presumably has some sort of appetite for food, but the total experience of such anticipation may be pleasant and even delicious, rather than unpleasant or painful. Similarly, in many other areas of human life. On this point see Kark Duncker’s “On Pleasure, Emotion, and Striving,” *Philosophy and Phenomenological Research* 1, 1941, esp. pp. 420–425. Life, human life, can be much better than the Buddhists (with their goal of nirvana) and the ancient Greeks (with their “soma sema,” i.e. the body is a

We have made this point, moreover, without bringing in a consideration that many ethicists would want to insist on, namely, that the worst people are incapable of the personal goods of friendship and love and so are lacking in happiness for such reasons. That might be true, but, then again, it might not be. Are malicious people incapable of having friends, are psychopaths incapable of friendship (e.g., as between psychopaths)? Was Stalin incapable of loving his mother or his daughter, Svetlana? Well, we don't have to investigate or pass judgment on such questions in order to make the points we are trying to make in this chapter. If the worst people are incapable of love and friendship, that adds fuel to the fire of our present argument that justice pays (more accurately, that the worst immorality doesn't pay). But whether they are or aren't, we still have the arguments that have been offered here, and I think they have some force toward showing, at long last and along psychological lines that have not previously been philosophically explored, that virtue and justice really do pay in self-interested terms.⁴

prison) realized; and it is worth noting that even when Buddhism was having its greatest influence on Chinese thought and culture, the Chinese were never convinced by this sort of pessimistic philosophy.

⁴There are two ways of interpreting the idea that justice pays. The first is that just people are likely to have better lives than the morally worst kind(s) of individual; and that is how we have been understanding the idea that justice pays here in this chapter. But the idea could be interpreted more strongly to mean that being unjust or being the worst kind of person *causes* one to have a worse life. This is clearly what Plato thought, but we haven't argued here that it is the immorality of serial killers and the like that causes their unhappiness. On the present account, rather, their unhappiness in some ways sets in first and leads them to be the kind of people they end up being. So it is only the first and weaker idea that conveys my intended thesis that justice pays and that indicates *how* I believe justice pays.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

