

The Beloved Self

The Holy Grail of Moral Philosophy

Yet according to the imperfect way in which human affairs are conducted, a sensible knave, in particular incidents, may think that an act of iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy. That HONESTY IS THE BEST POLICY, may be a good general rule, but is liable to many exceptions; and he, it may perhaps be thought, conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions. I must confess that, if a man think that this reasoning much requires an answer, it would be a little difficult to find any which will to him appear satisfactory and convincing.

Hume *An Enquiry Concerning the Principles of Morals*, section IX

At the start of book two of the *Republic*, Glaucon tells a story. A shepherd, Gyges, is watching over his flock when a storm breaks and an earthquake opens the ground next to him. Looking around, he finds a giant corpse wearing a gold ring, which he promptly removes. Twisting the ring this way and that, Gyges finds that he can make himself invisible and appear again. He immediately works out how to use this to his advantage. He seeks out the king's wife and seduces her. With her help he murders her husband, Candaules, and takes the crown for himself.

Gyges has done wrong. But which of us, Glaucon asks, would do anything else if we had this kind of power? He finds the story very disturbing, because he thinks that we all *ought* to be just even in the unusual circumstances described, when it would be in our interests to do otherwise. And he asks Socrates to prove that he is right.

Socrates spends the rest of the *Republic* trying to explain why Gyges really ought not to steal, kill and commit adultery. He tries to show that someone who is unjust is always psychologically damaged, and so, despite appearances, it is not in Gyges' interests to use the ring for nefarious purposes. Readers of the *Republic* tend to find Glaucon's question more compelling than Socrates' reply, however, and moral philosophers ever since have struggled to improve on it. Henry Sidgwick, for example, tries to defend morality in his monumental work *The Methods of Ethics* but it ends, literally, with the word "failure". As Hume reports with characteristically elegant understatement: it is difficult to give an argument that a "sensible knave" has reason to do what is morally right that is at all convincing.

Should we take Glaucon's problem seriously, though? The story of Gyges might be entertaining but it makes no claims to realism. We know that in real life, wrongdoing has consequences; no one could expect to emulate Gyges and get away with it. People who care about their own interests should for the most part do what is morally right. And indeed they do. Even the sensible knave admits that honesty is usually a good policy.

Glaucon thinks the story is important, however, and he is right. What Gyges should do once he has put on the ring shows something very significant about morality: it shows whether there are reasons to do what is morally right even when it is in your interests to do otherwise. We all know that justice and morality often have good consequences for you and injustice and immorality bad ones. The ring of Gyges story separates out the acts from their normal consequences. If you should still do what is morally right nevertheless, clearly you have a reason to do what is right which is not based on your self-interest.

According to Glaucon – and to most of us – morality makes demands on us and its demands do not cease simply because it is in our interests to do otherwise: Gyges has reason not to kill, for example, even if it would be advantageous to him. We may call this a part of *common sense morality*. According to common sense morality, morality has *practical authority* – we have reasons to do what is morally right and not to do what is morally wrong, and not merely because it is usually in our interests to do so.

Moreover, according to common sense morality, responding to these moral reasons for action is very important. Genuine moral action is not just a matter of doing the right action; to act well, you need to do the right action for the right reasons. So when the sensible knave does what is honest because it is in his interests to do so, he is not acting well. Of course, we often have no idea of another person's reasons for action. As morality and self-interest often coincide, the sensible knave will act outwardly just like someone who acts for moral reasons. We may not be able to tell why he acts until morality and self-interest no longer coincide, and he follows self-interest. This is another reason why the story of Gyges is intriguing. Up until he finds the ring, Gyges may have lived a morally irreproachable life – as far as anyone else can tell. But once he can turn himself invisible and get away with any kind of immorality, he follows self-interest. This gives us a good reason to suspect that all along, he was not acting well. He might have been doing the right thing, but not for the right reasons. We can see that this could be true of anyone. They might be doing the right thing, but from the wrong reasons, acting on self-interest alone.

Of course, the sensible knave disagrees with all of this. He cannot find any reason to be honest when dishonesty will benefit him. He claims that he has reason to do what is in his interests and no other reasons for action. He is an *egoist*. Though many of us do not agree with the knave, we do feel the temptation of acting from self-interest. What can we say to an egoist, or even to ourselves, to prove that we have reason to resist the temptation to do wrong? Can we find an argument that will persuade him that he has reason to be moral? A compelling argument against egoism has been described as *the Holy Grail of moral philosophy*. Many moral philosophers feel themselves obliged to search, but the goal, like the grail, has proved strangely elusive.

How can you argue against egoism? You might try a similar tactic to Socrates, and attempt to show that doing wrong cannot really be in your interests after all. It might look as if Gyges profited from his crimes but if being moral is an essential component of a good life, then his life could hardly have gone worse. Aristotle thought that in order to lead a good life, one had to have the virtues, including justice and honesty, and modern virtue theorists who follow him tend to agree. But the sensible knave has an obvious riposte: their notion of the good life is bizarre. It is simply crazy to deny that Gyges had a good life: he married a queen! He ruled a kingdom! The knave can develop his own account of a good life and a corresponding theory of virtue, where the virtues are the traits of character that are most likely to allow you to end up with an attractive partner together with plenty of money and power. Now our question becomes: which of these conceptions of the good life and the virtues should we accept? Which do we have reason to try to achieve? We are more or less back where we started.

Kant took the challenge from egoism very seriously. He saw the appeal of self-interest, and indeed was worried that most of us succumbed to it most of the time. We might usually do the right thing, but only because it is in our interests to do so:

I am willing to admit - out of sheer generosity! - that most of our actions are in accord with duty; but if we look more closely at our thoughts and aspirations we keep encountering the beloved *self* as what our plans rely on, rather than the stern command of duty with its frequent calls for *self-denial*.¹

Kant thinks that morality has an especially great authority: when morality commands us, we *must* obey. So if we do act wrongly, our choice is not simply morally at fault. It is a choice we have an overriding reason not to make. Kant takes up the challenge of defending common sense morality by trying to show that a fully rational person would never act immorally.

Kant's response to the knave is hard to understand, however, partly because it involves perhaps the deepest and most difficult part of his philosophy, his theory of freedom; partly because his target is not very plain. When the knave puts his beloved self in place of

¹ Kant, *G. 4*: 407. This translation is by Jonathan Bennett from his website: <http://www.earlymoderntexts.com/>

stern duty, Kant seems to think that he is putting his happiness first. But since Kant is not very clear about what happiness or the pursuit of happiness is, it is difficult to be certain exactly what his objections to it are and hard to assess whether they succeed.

Moreover, just as the knave can answer the virtue theorist by producing his own account of the good life and the virtues, he can answer Kant by doing the same, devising a parallel to Kant's moral law. We can call this *Kantian egoism*. Once again, the challenge becomes to defend morality against the knave's rival account of reasons for action.

This book is about the quest for the Holy Grail of moral philosophy: the challenge from egoism and how we can respond to it. The first part of the book clarifies the challenge by looking in more detail at the opponent, the sensible knave, setting out a small but important selection from the wide variety of theories of egoism. It begins with the most familiar, standard egoism, the theory that each person has reason only to maximize his or her own happiness. In chapter two I compare standard egoism to utilitarianism, pointing out the obvious similarities between the two. Next, the focus turns to Kant, and what it might mean for one's plans to rely on the "beloved self": first, Kant's own idiosyncratic account of the pursuit of happiness is discussed, and second, the theory that parallels the moral law is laid out. In chapter four, the knave's version of virtue theory is explored.

The point of this part of the book is not to defend any particular egoist theory, or indeed to defend egoism generally – in fact the hope, later in the book, is to vindicate morality, not self-interest. Consequently, I make no attempt to assess each version, to determine which is the most plausible or whether any is acceptable. I do think, however, that reflection on the moral reasons for action that different moral theories propose can help us to develop novel and interesting conceptions of self-interest and of reasons for action based on self-interest, and this in turn can illuminate aspects of the original moral theory. For example, in chapter three, I discuss in some detail what it means for rational nature to be a source of value, and give a new account of this key Kantian idea. My purpose in Part One, however, is to prove that egoism is both more interesting and more various than it is sometimes regarded, for standard egoism is often the only version ever discussed. Whilst it may be the most popular and most traditional form of egoism, there are many others and it is much harder to argue against egoism if you have to take on all versions of the theory, not just one.

The remainder of the book takes up egoism's challenge to morality and the ways in which philosophers could and should have tried to answer it. Parts two and three of the book can to some extent be read independently of part one, but not completely: some of the arguments against egoism apply to all egoist theories, but some apply to only one or more versions of the theory and fail against others.

In chapter five, I distinguish some more or less ambitious ways in which we might try to respond to an egoist. The ideal, the Holy Grail, is bold: an argument that even an egoist should accept that egoism is false. Some philosophers including Parfit, Korsgaard and Nagel have tried to formulate arguments along these lines, casting doubt on whether the knave's idea of reasons for action or his conception of personal identity is satisfactory. It is widely accepted that these arguments do not work. I show that they do indeed fail, and give some reasons why others are not likely to succeed either. Egoism is not internally inconsistent, nor has any argument against it, based on premises that an egoist should accept, been successful. The prospects for an ambitious vindication of morality are bleak.

Does this matter? It is not unusual in philosophy to be unable to answer an opponent in her own terms. Consider a sceptic who thinks that we have no knowledge of the external world. It is difficult, and perhaps even impossible, to prove to such a sceptic that we do have such knowledge. It is easier – though hardly straightforward – to show to our own satisfaction that we do. This vindication of our knowledge of the external world may be modest, but it is nonetheless important. It looks as if the best response to the egoist might be similar. We might just insist that we do have moral reason to help others, not to lie, not to kill and so on – and since egoism denies this it must be false. Of course, no sensible knave is going to be very impressed by this argument. But now we have given up trying to answer him in his own terms, and we are searching for something less demanding: an argument that we at least find compelling.

A modest reply to an egoist is an argument based on premises that the egoist denies, but that we accept. This kind of argument looks like it should be easier to provide than a more ambitious argument trying to defeat the egoist in her own terms. But modest replies to egoism encounter severe problems. I consider two of them. First, in chapter 6, I defend these arguments against the charge that they are not “cogent”, that is, they cannot resolve doubt about their conclusions or convince anyone of their truth. I argue that this problem is not insurmountable: modest arguments against egoism can be cogent. Unfortunately, they encounter a more serious difficulty, the subject of chapter 7: the problem of disagreement. The very fact that there is considerable disagreement in ethics, between moral agents as well as with egoists, mean that it is exceptionally difficult to establish a justification for the claim that we have moral reasons to act. As Part Two closes, it looks as if we may not even be able to give a modest defence of morality.

The arguments of Part Two of the book are overwhelming negative, suggesting that attempts to defend egoism, either ambitious or modest, fail. But in the most important part of the book, Part Three, I turn to the positive project of defending morality. I show that it is reasonable for us to accept that we have moral reasons for action. So *we can modestly vindicate morality* after all. And at the same time I develop an extremely unusual objection to egoism: an *epistemic problem for egoism*. The argument of this part of the book is complicated and in the course of it, I will defend several extremely controversial claims about moral epistemology, the function of moral beliefs and the nature of moral action.

I argue that there is a significant difference between the standards of epistemic rationality for beliefs about explicitly moral matters (including moral reasons for action) and the standards for beliefs about non-moral factual matters. In chapter 8, I suggest that the standards of epistemic rationality are connected to the function that those beliefs will play, including their role in contributing to action. In the following two chapters, I argue that moral beliefs (including beliefs about moral reasons for action) play a different role in action than beliefs about non-moral matters, and show that it follows that different standards of epistemic rationality apply to each. The standards of epistemic rationality for ordinary non-moral beliefs are related to knowledge, whereas the standards for moral beliefs are related to something different: *moral understanding*. I distinguish moral knowledge and moral understanding in chapter 9 and argue that it is moral understanding, not moral knowledge, that is most important: moral understanding plays a crucial role in moral virtue and in morally worthy action. In doing so, I give a new account of what it is to do the right act for the right reasons.

In chapter 10 I explain which responses to moral testimony and moral disagreement are appropriate if you are aiming to acquire and use moral understanding rather than moral knowledge. In chapter 11 I draw out the implications of all of this for disagreements between supporters of common sense morality and egoists, and draw conclusions about the epistemic standing of common sense morality and egoism.

I show that there is a striking *epistemological asymmetry* between the most plausible forms of egoism and morality. In their own terms, it is not epistemically rational for those egoists to believe their own theory, but it is epistemically rational for moral agents to believe that morality has authority. It turns out that we can defend morality - at least to our own satisfaction - and it is egoists who are left struggling for a view that is both plausible and epistemically respectable. We may not have vanquished egoism entirely, but it is common sense morality that ends up in by far the stronger position.