

Pathways to Philosophy

PROGRAM E: MORAL PHILOSOPHY

Reason, Values and Conduct: Unit Two

(a) prudential reasoning and weakness of will

24. THE project of the moral objectivist is to demonstrate a rational basis for moral conduct. To say that moral considerations are ultimately objective, not subjective, is to claim that morality is in some way founded in the nature of reality, rather than in our own contingent desires or feelings. There are things that are desirable – actions to be done or not done, virtues to be aspired to, vices to be avoided, perhaps feelings to be felt or not felt – not because of the way we as individuals or as a social group subjectively happen to be but because they are demanded by reason itself. Nor do we have a choice whether or not to throw our lot in with rationality. For such a choice could only be motivated either by reason or by irrational desire. If it is motivated by mere desire, then it would be open to someone, faced with moral criticism, to say, ‘I choose not to be what you call “rational”,’ in which case there would be no arguing with them. If, on the other hand, the choice is motivated by reason, then it is no choice at all but merely the re-iteration of an existing commitment. In short, if, as the objectivist claims, rationally guided action is equivalent to moral rationality, then to be rational cannot be an option – at least for beings for whom the question, ‘Why should I be moral?’ arises in the first place – but only a necessity.

25. To get a feel for what is at stake here, let us put questions of morality on one side for the moment, and consider examples of the application of reason to decision making where moral right or wrong is not in question. The laws of logic constrain the things we can rationally desire, just as they constrain the things we can rationally believe. The child says, ‘I want X,’ and then a few

moments later, 'I want Y.' We point out that Y would be equivalent to not-X. But still the child persists. Finally, after much anguish and commotion, the point gets home. It is one of the early lessons in life: You can't have your cake and eat it; you can't have things both ways. In formal terms, the child has grasped an application of the logical law, 'Not (P and not-P)': the Law of Non- Contradiction. Of course, people are illogical at times. No-one is physically compelled to be rational so long as they retain their freedom of action. Sometimes a person has a blind spot about a particular subject, or finds it difficult to think a problem through, or see the whole picture. And then again there are occasions where after long deliberation one discovers, or alternatively fools oneself into thinking one has discovered, a way of getting Y while avoiding not-X. However, such apparent exceptions would not tempt the philosopher to argue that the laws of logic are subjective and not objective. The validity of logical laws does not presuppose that one has 'chosen to be logical'.

26. With prudential reasoning, on the other hand, we are already entangled in philosophical controversy. Is it necessarily irrational to do an action A now which you know for certain that you will regret later? Of course, if one of your firmly held principles is, 'Never to do anything which I shall later regret,' then the intention to do A, together with the premiss that one will later regret doing A, contradicts the intention never to do what one will later regret. But is it necessary to hold that general principle? Is someone who knowingly refuses to be bound by the principle, or at least bound by it in every case, necessarily being irrational? In *The Possibility of Altruism* (Oxford 1970) Thomas Nagel argues that failure to conform your intentions to what the future 'you' will want can be interpreted as a failure to *identify* with that future self. There does seem to be a problem making sense of the actions of someone who had no conception at all of a future life that they were moving towards, or that was moving towards them: a life that was their very own, and not simply the life of someone 'like' them in certain respects. Yet for all that we do sometimes feel quite justified in prefacing a decision with the words, 'I know I am going to regret this but...'. There are times when one applauds imprudent spontaneity. Indeed, on reflection, a life of unbending obligation to all of one's future selves is hardly an attractive prospect. In the bloom of youth, we are not bound always to respect the second thoughts of crusty old age. This is merely a reflection of the fact that prudential

considerations do not over-ride all other considerations: there are no demands of prudence comparable to the demands of morality (unit 1/10).

27. Nagel, meanwhile, goes on to argue by analogy with prudence that our failure to take other persons into account in our actions – to act ‘altruistically’, in the sense of giving the desires of the other equal weight to our own desires – may be interpreted as failure to acknowledge, or acknowledge fully the *reality* of another mind, the refusal to recognise the other as fully equal to myself from the objective standpoint. Once again, however, the question arises whether, if this is indeed the case, reason demands that one ‘believe in other minds’, why adopting the goals of the psychopath is necessarily a case of irrationality. And even if one does ‘believe’, no reason has been given why that should not be compatible with a perverted ‘acknowledgement’ of the other, in the way the torturer gains pleasure from appreciating the very real pain of his victim. (The analogy with prudence here might be ‘cutting off one’s nose to spite one’s face’: far from not minding the regrets that will be voiced by my future self, I act deliberately in spite of them. Full of remorse, say, at my cowardly decision to agree to do Y, I knowingly throw a spanner in the works, spoiling in advance any chance of success and savagely punishing my future self into the bargain. I do this precisely because I recognise that that future self is me.)

28. A far more radical objection to Nagel’s argument comes from Derek Parfit in his book *Reasons and Persons* (Oxford 1984). According to Parfit, once one discards the Cartesian belief in an indestructible, immutable soul – and with it the idea of the pre-eminent importance of personal identity – my future self has to be regarded as merely one person amongst many other persons who are more or less ‘like’ my present self; all of whose desires I ought rationally to grant equal weight to my own desires. Thus, a wonder drug might conceivably be discovered that prolonged human life by thousands of years, over which time I could gradually lose every trace of memory and psychological continuity that connected me to ‘my’ former selves. Or a machine for physically duplicating human beings might be constructed: I step into the booth, and a minute later out come two physically and mentally indistinguishable versions of ‘me’. What these thought experiments show is that, barring belief in a Cartesian soul, the relation we term ‘personal identity’ cannot, in principle, be all-or-nothing or even one-to-

one. What makes me 'me' is simply a structure of a certain kind – say, the program physically encoded in my brain – and every structure can, in principle, be duplicated, or change by slight degrees into an altogether different structure. The lack of an ultimate basis for the rationality of prudence, as Nagel interprets that notion, turns out for Parfit to be the very principle behind the rationality of moral conduct.

29. The question raised by Parfit concerning the importance of personal identity, with its implied attack on the special value of personal integrity, is one to which we shall return when we consider accounts of moral conduct based on the utilitarian principle of the 'greatest happiness for the greatest number'. Whether or not Parfit is right in criticising the importance of personal identity, however, prudential reasoning undoubtedly plays a significant and valid role in our everyday deliberations. It therefore provides a neutral testing ground for two issues which have an important impact on a philosophical account of moral decision making, namely, weakness of the will (1/19) and freedom of the will (1/20-21). The way these problems are resolved in the case of prudential reasoning will provide valuable material for constructing our account of moral reasoning.

30. Sometimes we fail to do things that it would be best for us to do, and sometimes we do things that it would be best for us not to do. Moreover, we cannot always blame our regrettable action or inaction on a mere physical inability to carry our best counsel into effect, or on ignorance of the consequences of a decision made according to our best lights. On at least some occasions, it seems, we deliberately and knowingly choose the option which is worse for us, not out of a masochistic desire for self harm, but simply because our recognition of what we ought to do or not do in our own best interest is overruled by a more immediate aversion or temptation. Not everyone appears equal in this respect. We admire some persons for their self control or 'will power' while despising others for their incontinence. These character traits, it should be noted, are not in themselves deserving of moral praise or censure. The 'strength of will' of the individual ruthlessly pursuing his or her self-interest to the exclusion of all else makes them, in our eyes, all the more evil and despicable. (On a popular misreading of Nietzsche, all values are ultimately reducible to the scale of relative 'will to power'. Later in the programme, when we come to consider the question the question of how

values arise – the extent to which they are ‘created’ or ‘discovered’ – we shall see that this was far from being Nietzsche’s considered view.) It seems highly likely, in any event, that the analysis of the formal structure of weakness of the will in the case of imprudence will carry over with little or no modification to moral reasoning.

31. The first picture that comes to mind when one thinks of weakness of will is that of a human being pushed and pulled by various internal forces. Some of these forces we call ‘impulses’ or ‘emotions’, ‘temptations’ or ‘aversions’. Others appear as commands. When we act according to reason, we obey a command that comes from within, doing what we ourselves recognise we ‘ought’ to do. In acting against reason, on the other hand, we choose to disobey this internal voice. The force of temptation or aversion proves stronger than the capacity of the internal voice to inspire obedience. In the moral sphere, the internal voice is called ‘conscience’. (There appears to be no corresponding term in the case of prudential reasoning. We do not talk of having a ‘bad conscience’ about our failure, say, to follow doctor’s orders – except in an implicitly moral sense when we feel guilty at letting down someone who has our best interests at heart.) At any rate, the popularity of this view of human deliberation is reflected in the language we use to describe such cases: ‘I told myself, “Come on now, you’ve got to get a grip on yourself!”, but finally I just had to let go.’ ‘I felt my resolve slipping away, but I took control of myself at the last minute.’ – How coherent is this picture?

32. The idea of the self split into different parts traces back to Plato, in his famous image of the charioteer of reason battling against our animal instincts to keep control of the unruly steeds of the will. In our capacity to act and initiate changes in the physical world we are like sources of energy capable of being turned in different directions. The control of this energy, meanwhile, falls between our physical nature on the one hand and a force arising from outside the physical realm, the impulse that comes from our rational nature or non-physical ‘soul’. In non-human animals, by contrast, the only conflict that can take place arises between conflicting physical impulses: for example, the dog’s desire for a juicy steak versus its fear of punishment if it attempts to remove food from the dining room table. Grasping the steak does not indicate weakness of will on the dog’s part, but merely a lack of sufficient training. In a conflict between purely physical impulses, the mathematically strongest

desire or combination of desires wins out just as surely as the motion of an object according to Newton can be predicted by calculating the net result of the forces acting upon it.

33. Transposed to the human sphere, the idea of human action as the net result of internal pushes and pulls over which the agent has no ultimate control appears the very picture of unfreedom. All deliberation concerning what we 'ought' or 'ought not' to do is a sham. Or, at least, what we term 'deliberation' is in reality either the mere calculation of consequences, or else a device for self-seduction: I make the prospect of the doctor's nasty medicine appear more desirable by conjuring up in my mind images of my ravaged state of health if I continue to refuse to take it. Yet many would find that picture vastly preferable to the Platonic alternative where human beings are split down the middle between a rational side and a non-rational side, with no intelligible connection between the two halves thus separated. The eighteenth century German philosopher Immanuel Kant later took up the Platonic view, rendering its metaphysical assumptions all-too apparent. According to his theory of 'transcendental idealism', there are two 'selves' or 'I's' belonging to two separate 'worlds': the phenomenal world in space and time, where all movement and change is governed by the law of cause and effect, and the noumenal world of 'things in themselves' that belong neither in space nor in time. It is my noumenal self that acts according to laws of practical reason – which Kant equated with the laws of morality – and in so doing demonstrates its ultimate, metaphysical freedom. Yet Kant gave no clear explanation of how the noumenal self and the empirical self are able to battle it out for supremacy in cases where the agent is torn between doing what they ought to do and what they want to do, nor, given that our bodily movements – including the 'explanations' we offer of our actions in speech or writing – are wholly explicable in terms of causal laws, how the 'decisions' issuing from our noumenal self can have any significance whatsoever. The most it seems Kant is entitled to is the tautological claim that it is only insofar as our physical actions conform to the 'laws of rationality' posited by the logician or the moral philosopher that they are 'rationally intelligible'.

34. However, there is a third alternative. We are, at one and the same time, animals motivated by bodily needs and desires, and – unique in the animal kingdom – rational agents and language users. However, to say that we are

rational is not to say that we have a rational 'side', in addition to a non-rational 'side' but rather that all our decisions are governed by reasons of one sort or another: we are 'rational' through and through. That is true even when we fail to act according to the best reasons: only a rational agent can act irrationally. As the philosopher Elizabeth Anscombe notes in her book *Intention* (Oxford 1972), the statement, 'I want X' requires a 'desirability characterisation' in order to make it intelligible as an explanation of action: a concept under which X falls that gives a reason for wanting X, a reason that would be recognised by other language users, however much one might disapprove of, or regard as foolish actions done for that reason or that reason alone. To quote Anscombe's example, we would simply not be able to make sense of an individual's declared desire for a 'saucer of mud', unless they went on to explain what they wanted to do with it, or what it would do for them. ('I need it to feed my algae.' 'I just love the smell.') Thus, one might argue, the most peculiar food fad is recognisable as an explanation of human action only insofar as the 'desire for X' falls under concepts, say, of taste or nourishment. The most exotic sexual preference is describable as such only insofar as one can represent, in the agent's eyes – however unlikely it might seem to those who regard themselves as 'normal' – the desirability of doing Y, or doing Y to Z. (In the case of actions or motivations that appear on the surface simply unintelligible, a lot of work may be required on the part of a psychotherapist, say, in helping the agent bring the underlying reasons to light.)

35. If we are 'rational through and through', and not beings who are part rational and part non-rational, then cases of weakness of will are necessarily cases of failure of reasoning. In weighing up conflicting reasons for or against a given action, we appear paradoxically to recognise one side of the argument as having the 'best' case, while allowing the 'worse' case to win the argument. It is as if knowledge speaks quietly, while error shouts. Yet that is hardly in itself sufficient explanation. If we do 'hear' knowledge speak, if we recognise which reasons are strongest, then however quietly those reasons speak we ought, logically, to act accordingly. That is just what recognising a reason as 'the best' or 'the strongest' means. So what is it that goes wrong? Consider again the example of the nasty medicine. (Just to make the example more realistic, it might be a particularly unpleasant course of chemotherapy for a cancer; or perhaps the amputation of a gangrened limb.) Let us assume that

there is no reasonable question of weighing up the cost of 'taking the medicine' in relation to its accepted benefits. There is no question but that I want to live. But now I walk into the cancer ward and see for the first time the ashen faces and hairless heads; or I have just watched the soldier in the next bed have his leg sawn off without anaesthetic, and it's my turn. The slightest reasons for doubt over the necessity of the prescribed course of action are magnified a thousand-fold. Cowardice can make the best of us willing victims of self-deception. (Some might consider that it was better to die rather than face terrible pain, even if they were fully convinced that the pain will pass; but we are not considering that case.)

36. It is a form of self-induced blindness, on this account of the formal structure of weakness of will, that leads me, say, to accept the third whisky, even though I know I ought not to drink and drive, or to light up another cigarette in spite of my chronic bronchitis. Simply citing 'cowardice' in facing up to doing what has to be done, or 'weakness' in the face of temptation by itself explains nothing. The fear, the temptation, are as such perfectly reasonable. Other things being equal, one has every reason to pursue what one perceives as pleasure and avoid what one perceives as pain. Yet in the examples we are considering 'other things' are not equal. The reasons for pursuing pleasure or avoiding pain in these cases are less strong than other, countervailing reasons. The only possible explanation for our failure to act accordingly, on the model of choosing between competing reasons, is that at a critical moment we become temporarily blinded. Our failure is a failure of knowledge. How is that possible? – One can make a start on that question by pointing out that certain mental habits can strengthen or weaken our capacity for keeping our eyes fixed on the reasons that appear to our considered judgement the strongest. That is not necessarily a reflection of our intellectual ability to follow a chain of reasoning, nor of the strength of any given belief or desire, but rather a disposition that we have acquired, over a period of time, to conduct our practical deliberations in a certain *manner* or frame of mind: to be constant or inconstant, focused or unfocused, capable of detachment or unable to rise above the heat of the moment. It is just as much a trait of character as decisiveness or indecision, industry or laziness, temperance or intemperance.

37. In the light of this, 'weakness of will' would appear a misnomer. The will

as such is neither strong nor weak; it is not a source of raw energy or power that some have more of than others (though undoubtedly some persons have more mental or physical energy than others). The higher a given objective appears on your scale of preferences, the more vigorously you will pursue it – other things being equal. But something else is required to avoid succumbing in the face of contrary ‘temptations’; and that is the ability to maintain an unclouded perception of that objective when the going gets tough. That is a trait of the intellect rather than the will, cognitive rather than affective. Weakness of will is, in short, a species of irrationality.

38. In these terms, one can construct a partial vindication of Kant’s view. There are indeed, as Kant believed, two fundamentally different vantage points that one can take up in explaining someone’s actions. But these two standpoints are not, as Kant thought, explanation in terms of the subject’s reasons and explanation in terms of the subject’s ‘irrational’ desires. Rather, we can draw back from considering the reasons for action that appear to the perspective of the agent and focus instead on the degree to which the agent’s general behaviour exemplifies certain laws or ideals of rationality; the underlying explanation of the agent’s failure to correspond to the ideal lying not in the domain of reasons but rather of causes and effects. As ‘reasoning machines’, some individuals are constructed better than others (just as some calculating machines or computer chips are better built or designed than others). Some of us, however, improve with time (an improvement, it should be noted, that we can rationally aim to bring about by cultivating better mental habits through training and practice) while others, sadly, get worse.

(b) weakness of will and moral reasoning

39. All we have said about the formal structure of weakness of will applies with equal force to prudential and to moral reasoning. However, in the case of moral reasoning the stakes are much higher. Prudential reasons can be rationally over-ridden while moral reasons cannot (26). If sensitivity to moral considerations is a matter of objective knowledge rather than subjective attitudes, then our failure to do the right thing must always be due to a failure of knowledge. (The only alternative, as we have seen, leaves the self split in two, with no intelligible account of how the two sides connect with one another, 33.) That is what Socrates taught. The moral virtues – such as justice, courage, temperance – are all species of knowledge. (Hence the other famous Socratic doctrine: the ‘unity’ of the virtues.) Learning to philosophise about the nature of the virtues by means of the open-ended, question-and-answer method of Socratic dialogue makes the student morally better by increasing and developing their understanding of ethical matters. Yet such knowledge is as much a practical ability as a proficiency in dialectic. It is not enough to be able to come up with the right verbal responses: one has to learn to see the reality that the words are about. Otherwise, a view that one has been persuaded to accept, one can later just as easily be persuaded to reject. – By the same token, Socrates thought, a good man who ‘sees’ what is to be done but lacks philosophical understanding of the nature of the virtues is vulnerable to sophistical arguments that cast doubt upon the validity of his unreflective moral judgements.

40. The doctrine that virtue is knowledge is indeed paradoxical. Temptation or cowardice do not immediately present themselves as species of irrationality or of lack of knowledge but rather as mental disabilities that directly affect our capacity to put our considered decisions into effect. But that is merely part of the self-deception. The story we tell ourselves is one of mastery and obedience, of refusing to act as the voice of conscience dictates. As we stand, paralysed, unable to act, or alternatively feel our internal resistance to the prohibited deed crumbling away, self-exhortation seems powerless in the face of ‘the way I am’. Yet if weakness of will really were a species of disobedience, what are the sanctions? If there is no punishment for disobedience or reward for obedience, then why should I ever obey the voice

of my conscience? 'Because it is mine,' is no answer. By talking of one who commands and one who is commanded, I have in effect made the voice of 'my' conscience something other than myself, the subject. (According to Freud, the voice of conscience is the superego, the internalised parent.) Nor does it help to point out that the feeling of guilt is a kind of punishment, and the feeling of self-satisfaction a kind of reward. We do not 'obey our conscience' simply in order to avoid guilt or obtain self-satisfaction. It would be a totally unconvincing diagnosis of weakness of will to posit that in such cases what has gone wrong is simply that the internal pricks of conscience are not sufficiently penetrating, or that the glow of self-satisfaction is not sufficiently warming. If that point is not obvious, one can gain further confirmation by turning back to the case of prudential deliberation, where talk of 'conscience' is redundant (31) and reward and punishment are built in to the set up. If the prospect of a very real reward or punishment is not enough to make me act in what I judge to be my best self-interest then some other account must be given of my failure to do what I believe I 'ought'.

41. What then, finally, of our highly valued faculty of freedom of the will? Many would find the claim that 'freedom from determination' is not something to be desired (1/20-21) no less paradoxical than the claim that there is no such thing as a weak 'will'. Yet if the moral philosopher were to succeed in showing that right action is a matter of objective knowledge, and wrong action solely the result of ignorance or irrationality, then the only 'freedom' we should desire is the freedom to think things through for ourselves rather than be coerced by others; so that our actions are determined by our moral knowledge and powers of reasoning. By contrast, according to the Platonic or Kantian picture of a self split into non-rational and rational halves where the agent is required to 'choose' between following reason and acting in accordance with their 'animal' nature (33), 'freedom from determination' there undoubtedly is, but it is the freedom of an inexplicable leap in the dark; freedom in name only that merely serves to patch together an unintelligible theory. On the theory that views the self as fundamentally a unity, as 'rational through and through' (34), we are not, and should not wish to be 'free' to choose whether or not to act according to reasons. We are not, ultimately, 'free' to reason well or badly, or to choose whether or not we wish to know. At every moment, we have no choice but to act according to the way we see things, that is to say, the way the facts – the empirical facts or the

moral facts – impress themselves upon us. Whenever an agent’s moral perception is clouded, or they fail to think things through, wrong action may, and perhaps almost certainly will result. According to the objective view of moral reasoning, however, the remedy is not retribution but instruction.

42. Once more, we can gain illumination by looking at how the question of freedom arises in the case of prudential reasoning. When we blame ourselves for a wrong decision that adversely affects our own interests what exactly are we doing? We may regret not knowing then what we know now; with the benefit of hindsight we should not have acted as we did. We can go through the motions of our process of coming to a decision – or conjure up the state of mind that produced the instant response that seemed appropriate at the time – in just the same way as the tennis player who misses an easy return mimes the stroke they should have played as they prepare to face another serve, with the intention of ‘doing it right next time’. What room, one might ask, is there for the thought ‘I might have done otherwise than I did’? Given what I did know, given the way I did reason, I had to make the decision that I made. The way I saw things at the time – and of course the things I wanted at the time – ruled out any course of action other than the one I actually followed (assuming I did not act on the spin of a coin). If things had been different, I might have acted differently. Howsoever I did act, or howsoever I might have acted, there is simply no room for any notion of ‘freedom’ that implies that in the very same circumstances and the very same state of knowledge I might have acted differently.

43. According to the objective view of morals, knowledge, in the fullest sense, of what is morally right or wrong – of the ‘moral facts’ – is not only inseparable from right action, but fully determines right action. The will of the virtuous person is nothing but the instrument of their moral knowledge, the physical impulse to action arising from accurate moral perception and reasoning. (By complete contrast, the moral subjectivist must say – along with the philosopher David Hume – that reason ‘is, and ought only to be the slave of the passions’: *Treatise of Human Nature*, Book II, Part 3, Sec. 3.) What clue does this result give to whatever ‘strange kinds of objects’ (1/19) moral facts might be? It seems clear enough that moral facts are not ‘objects’ in any recognisable sense: neither physical objects nor metaphysical objects. A ‘physical moral object’ might be, say, a gene that innately predetermines the

disposition to morally right action. But supposing that such a gene were identified (and there is no objection in principle to such a contingency), those who discovered that they possessed the moral gene would not be under any obligation to obey the impulses generated by it. A 'metaphysical moral object', on the other hand, might embody a compelling representation of the good life or of moral virtues to be emulated. (That is how Plato conceived of the 'Forms' of Justice, Courage and the rest.) But whether or not a given representation is found compelling depends on the percipient. There is no necessity that it should inspire right action, or indeed why it should not inspire the impulse to do the very opposite of the actions it portrays in a supposedly favourable light. If there are moral facts, it is far more likely that they will be found in structural features that universally characterise the very process of practical reasoning. The 'moral' facts are simply the facts as such taken in the only way a moral being can take them. – Why we are necessarily 'moral beings' (including those individuals who think and act as if they were not) is the question that has yet to be answered.