

Three Methods and a Dualism

JOHN SKORUPSKI

SIDGWICK DISTINGUISHES three methods of ethics: intuitionism, egoism and impartialism (as I shall call it).¹ Yet he holds that just two of these, egoism and impartialism, are grounded in 'practical reason'—and in fact he famously concludes that these two methods constitute an irreducible 'dualism of the practical reason'.

We still remain, I think, uncertain of the relations between morality, self-interest, and an impartial theory of the good. And the persistence, integrity and penetration of Sidgwick's ethical thought give his conclusions as to these relations a continuing authority. However, they also raise questions in a usefully clear way. Notably: if there are three methods, why should there be only a *dualism* of practical reason? Or,

¹ Sidgwick calls it the method of utilitarianism—which he also calls 'universalistic hedonism'. However, the contrast that is of interest here (and which interested Sidgwick) is between the egoistic thesis that the good of any individual has an agent-relative rational claim on that individual's deliberation—together with the method(s) founded upon that view—and the impartial thesis that the good of any individual has an agent-neutral claim on anyone's deliberation, together with the method(s) founded on *that* view. For this purpose we do not need to decide on the truth or otherwise of hedonism, but neither do we need to decide on the specific distributive doctrines of aggregative or average utilitarianism. Impartialism as such says only that the good of any individual is agent-neutrally good, that is, that any individual has reason to promote it. Aggregative or average utilitarianism makes further assumptions about the form of the function from individuals' goods to the agent-neutral good, and there are many other plausible options. But the choice between them is not relevant here. See John Skorupski, *Ethical Explorations* (Oxford and New York: Oxford University Press, 1999), chs 3 and 5.

otherwise put, if the method of intuitionism is not grounded in practical reason, what establishes, of either of the other two, that *it is*? My conclusions will be anti-Sidgwickian. I shall argue that if 'practical reason' is construed broadly it is characterised by a wide diversity of irreducibly warranted—though not indefeasible—practical dictates or principles, of which prudence is one. On the other hand if practical reason is construed narrowly, conceived one might say as *pure* practical reason, then it exhibits a characteristic unity: so conceived, its sole warranted and indefeasible principle is a principle of impartiality in assessment of the good. Either way practical reason is not dual. These conclusions, as I say, are anti-Sidgwickian; I hope none the less that what follows can be accepted as a homage to Sidgwick.

1. *Intuition and reason*

The method of the intuitionist is that of investigating common-sense morality to identify its dictates in general or particular cases. Sidgwick calls these dictates 'intuitions', by which he means "immediate judgement[s] as to what ought to be done or aimed at".² They are 'immediate', he explains, in that they appear as *knowledge* in their own right and not as knowledge derived from something else. And he notes that any method of ethics must ultimately rest on at least one such 'intuition'—on a judgement as to what ought be done or aimed at which appears as evident in its own right; and which is, on that basis, regarded as 'immediately known to be true'.³

So when Sidgwick calls something an intuition he means that it appears as knowledge in its own right. He does not mean that it *is* knowledge. He makes this explicit later, when he notes that the term 'intuition'

has sometimes been understood to imply that the judgement or apparent perception so designated is *true*. I wish therefore to say expressly, that by calling any affirmation as to the rightness or wrongness of actions

² HS, *The Methods of Ethics*, 6th edn (London: Macmillan, 1901), 97.

³ In that sense, he notes, the egoistic and universalistic methods are also 'intuitional'.

'intuitive,' I do not mean to prejudge the question as to its ultimate validity, when philosophically considered: I only mean that its truth is apparently known immediately, and not as the result of reasoning. I admit the possibility that any such 'intuition' may turn out to have an element of error ... indeed the sequel will show that I hold this to be to an important extent the case with moral intuitions commonly so-called.⁴

What distinguishes the intuitionist is thus not just the claim that common-sense morality contains a large diversity of such intuitions, but the further claim that these intuitions, or at least some of them or some part of them, are authentic, underived bits of knowledge.

Does Sidgwick simply disagree with this claim? For a writer who makes such efforts at precision, I find it extraordinarily hard to be sure. One possible interpretation would say that he does. On this interpretation Sidgwick thinks that no principle of ordinary morality is *immediately known* to be true; such principles, or rather their more precise correlates, can only be *known* by being derived from the utilitarian principle which is itself immediately known. Thus the method of the intuitionist⁵ founders because it finds no genuine intuitive knowledge. On this interpretation it is not puzzling that Sidgwick thinks that there is only a dualism of practical reason even though there are three methods of ethics. For one of these methods turns out to yield no normative knowledge in its own right.

However, on another possible interpretation Sidgwick does not deny that common-sense morality contains some intuitive knowledge. On this interpretation, Sidgwick is prepared to accept that common moral intuition does yield some underived or immediate *knowledge* of normative truths, insisting only that its knowledge is unclear and imprecise—and indeed that it also contains an 'element of error'. On this interpretation his claim is only that the knowledge it does yield cannot be made 'scientific'—clear and precise—by the method of the intuitionist alone. Thus in the last paragraph of his 'Review of Common Sense' (*Methods*, Book III, ch. xi) Sidgwick says:

⁴ *Methods*, 211.

⁵ The 'dogmatic intuitionist': *ibid.*, 102. For discussion of how Sidgwick understands the term 'intuition', and of when intuition yields knowledge, I am indebted to Robert Shaver.

Nothing that I have said even tends to show that we have not distinct moral impulses, claiming authority over all others, and prescribing or forbidding kinds of conduct as to which there is a rough general agreement, at least among educated persons of the same age and country. It is only maintained that the objects of these impulses do not admit of being scientifically determined by any reflective analysis of common sense. ... the Morality of Common Sense may still be perfectly adequate to give practical guidance to common people in common circumstances: but the attempt to elevate it into a system of Intuitional Ethics brings its inevitable imperfections into prominence without helping us to remove them.⁶

Here Sidgwick does not, it is true, explicitly say that these 'distinct moral impulses' amount to intuitive knowledge—only that they are 'perfectly adequate to give practical guidance'. But in what sense 'adequate'? The most obvious reading is that they yield adequate *knowledge* for practical purposes. For Sidgwick's only point against them is that their objects cannot be 'scientifically determined'—made precise—by reflective common sense alone. And it is obvious that there can be knowledge that is not *precise* knowledge, so unless Sidgwick has surprisingly overlooked this point he could perfectly well grant that common-sense morality yields such knowledge, which is adequate for (most) practical purposes.

Consider how he deals with a particular such moral impulse, that of gratitude:

the duty of requiting benefits seems to be recognized wherever morality extends; and Intuitionists have justly pointed to this recognition as an instance of a universal intuition. Still, though the general force of the obligation is *not open to doubt (except of a sweeping and abstract kind with which we have not here to deal)* its nature and extent are by no means equally clear.⁷

He proceeds to highlight the unclarities. But pointing out such unclarities is perfectly consistent with accepting that I know the following: (a) that if a person has done me a lot of good out of sheer good will, I have reason to be grateful and show it; (b) that lack of gratitude, taken far

⁶ *Methods*, 360–61.

⁷ *Ibid.*, 259–60, my emphasis. I take it that the sweeping and abstract kind of doubt could also be raised about my knowledge that there is a desk in front of me.

enough, can become blameworthy. The unclarity, or rather imprecision, of (a) and (b) is evident, but so is their truth. Would Sidgwick, if challenged, deny that we '*immediately*' know that truth? On the second interpretation, he would not. On this interpretation Sidgwick's view is that common-sense morality contains intuitive knowledge, though through lack of precision it cannot qualify as 'scientific' or 'philosophical' knowledge.⁸

In which case, why does Sidgwick think there is only a dualism of practical reason? Why doesn't the intuitive knowledge found by the intuitionistic method also count as part of practical reason? Is Sidgwick just identifying—indeed confusing—rationality with clarity and precision? Why should purely *rational* intuitions have to be clear and precise in a way that (a) and (b) are not? It would require substantive philosophical argument to establish that. Anyone who merely assumes it evinces a preconception about practical reason which could fairly be called rationalistic or scientific.

I have just used the phrase 'purely rational intuition'. And this raises a variety of further questions. Is there a difference between an intuition and a purely rational intuition? What are we to understand by practical reason anyway? Further, if there is a plausible account of practical reason, or as I shall suggest later, *pure* practical reason, on

⁸ Bk III, ch. xi sets up four criteria for 'moral axioms' (see §2, *ibid.*, 338–43). They should (1) be stated in 'clear and precise terms', (2) be 'really self-evident', (3) not conflict 'with any other truth', and (4) be supported by an 'adequate "consensus of experts"'. He claims that no ordinary moral principles can satisfy (1). But his own list seems to allow that they could fail (1) and satisfy (2)–(4)—in particular, they could be imprecise and yet still 'really self-evident'.

Further, in his chapter on 'The Relation of Utilitarianism to the Morality of Common Sense' (Bk IV, ch. iii) Sidgwick explicitly disclaims the hypothesis that 'the perception of the rightness of any kind of conduct has always—or even ordinarily—been derived by conscious inference from a perception of consequent advantages'. *Ibid.*, 457. It is, he concedes, 'not as the mode of regulating conduct with which mankind began, but rather as that to which we can now see that human development has been always tending, as the adult and not the germinal form of Morality, that Utilitarianism may most reasonably claim the acceptance of Common Sense'. *Ibid.*, 457. This raises very interesting questions about the connections between morality's epistemology and its history. And it's at least compatible with the (second interpretation) view that the 'intuitions' of common-sense morality have an immediate or underived default warrant, and at least in some circumstances constitute knowledge.

which the immediate intuitions of common-sense morality don't belong to pure practical reason, what epistemic status do they have and in particular how do they constitute knowledge?

We shall not be able to address the last of these questions at all fully. However, I shall argue that a distinction between 'intuition' and 'purely rational intuition' is not pointless. Many ordinary moral judgements are warranted in their own right without being derived, directly or indirectly, from any other more basic normative principle; in an important sense however they are not judgements of pure practical reason. The judgement that someone has acted in a blameworthy ungrateful way would be one example. On the second interpretation Sidgwick could agree with this. Contrary to Sidgwick, however, I shall argue that the egoist's principle must also be seen, epistemologically, in much the same way as these ordinary moral 'intuitions': like them, it is derived from nothing more basic—but like them, it is not a principle of pure practical reason either.

Distinct and irreducible sources of practical reasons, I shall suggest, underlie *each* of Sidgwick's three methods. But only the impartialist's principle, the principle that the good or well-being of any being is, simply, good, has its source in *pure* practical reason. In contrast (I shall argue), neither the reason-giving force of prudence nor that of morality is reducible to pure practical reason. In both cases it derives instead from the hermeneutics of the sentiments: in the case of morality, the sentiment of blame; in the case of prudence, that of desire.

To make this clear we shall have to distinguish carefully two separate questions. The first concerns the epistemic pedigree of a practical principle. Here the question is, from what source does the principle derive its (default) warrant? The second concerns its potential defeaters. The question in this case is whether the principle can be defeated by different principles stemming from another source.⁹

⁹ I have shifted here from the concept of knowledge to that of warrant. Note that (i) a belief may be warranted but not true; (ii) a warranted belief may cease to be warranted in an improved state of information. The question Sidgwick raises about common moral intuition is whether it delivers any immediate knowledge. It could also be asked whether common moral intuition delivers any immediately warranted beliefs. I return to these questions in sections 5 and 6.

But before moving to these claims we must consider a perplexing question: how to *formulate* the alleged dualism of practical reason.

2. *The dualism of practical reason*

The preface to the posthumous sixth edition of *The Methods of Ethics* prints a manuscript draft in which Sidgwick gives a brief account of the development of his ethical view.¹⁰

Beginning, he writes, from an adherence to Mill's utilitarianism, he came to worry that it did not deal adequately with the conflict between self-interest and duty. He reread Kant and was impressed by Kant's fundamental principle, which he formulated for himself as the principle 'That whatever is right for me must be right for all persons in similar circumstances'. This, he thought, was 'certainly fundamental, certainly true, and not without practical importance'.¹¹ But it did not meet the difficulty which had led him from Mill to Kant: 'it did not settle finally the subordination of Self-Interest to Duty'. For a rational egoist could accept it. As a *rational* egoist he would accept that when any person is faced with a choice between his own and the general happiness the right thing for that person to do is to choose his own. 'The rationality of self-regard', Sidgwick continues, 'seemed to me as undeniable as the rationality of self-sacrifice. I could not give up this conviction, though neither of my masters, neither Kant nor Mill, seemed willing to admit it: in different ways, each in his own way, they refused to admit it.'¹² Delving back further in the history of ethics, he found that Butler had affirmed the same duality of interest and duty: 'he recognised a "Dualism of the Governing Faculty"—or as I prefer to say "Dualism of the Practical Reason," since the "authority" on which Butler laid stress must present itself to my mind as the authority of reason, before I can admit it'.¹³ The final step was provided by Aristotle. Sidgwick came to see the significance of

¹⁰ *Methods*, xv–xxi.

¹¹ *Ibid.*, xvii.

¹² *Ibid.*, xviii.

¹³ *Ibid.*, xix.

Aristotle's interrogation of common-sense morality and the need to do the same himself.

But the result of this examination was to bring out with fresh force and vividness the difference between the maxims of Common Sense Morality (even the strongest and strictest, *e.g.* Veracity and Good Faith) and the intuitions which I had already attained, *i.e.* the Kantian Principle ... and the Fundamental Principle of Utilitarianism.¹⁴

He became increasingly convinced that common-sense morality was a 'system of rules tending to the promotion of general happiness. ... the morality of common sense showed me no clear and self-evident principles except such as were perfectly consistent with Utilitarianism'.¹⁵

Two points are noteworthy. First, in this account Sidgwick does not state any definite principle of rational egoism.¹⁶ He merely points out egoism's consistency with the Kantian principle, which he understands as a principle of universalisability. Second, the answer to the question with which we were concerned in section 1, the question of what his view of the standing of common-sense moral maxims is, remains unclear here too. He says that his investigation forcefully and vividly showed the difference between them and the Kantian and utilitarian principles. He also says that the only moral maxims which seemed clear and self-evident were such as were consistent with the utilitarian—or in our terms, impartial—principle. Is he then accepting that some principles of the morality of common sense *are* clear and self-evident? On this reading (the second interpretation) ordinary morality contains principles which are self-evident in their own right—but they turn out on examination not to conflict with the impartial principle, whereas egoism and impartialism, also separately self-evident, do appear to conflict (unless we assume a moral government of the

¹⁴ *Methods*, xx.

¹⁵ *Ibid.*, xxi.

¹⁶ Indeed he never provides a suitable formulation of the egoist principle. Most of his formulations are statements of the irrationality of pure time-preference in decisions concerning one's own good—to that extent they are not clearly at odds with the impartialist's principle at all. Jerome B. Schneewind, *Sidgwick's Ethics and Victorian Moral Philosophy* (Oxford: Clarendon Press, 1977), chs 10 and 13, provides a valuable survey and discussion of Sidgwick's various formulations of his principles or axioms.

universe to reconcile them). And this then explains why Sidgwick is concerned with that particular dualism.

Clearly Sidgwick does not regard the dualism as an outright contradiction. For his point is that the two principles may come into conflict in the absence of a moral government of the universe. So the principles of egoism and impartialism should not be formulated in such a way as to be inconsistent outright. How then should they be formulated?

One might first of all consider the following:

- (i) The degree to which there is reason for me to do an action is proportional to the degree to which it promotes the good of beings overall, taking the good of all beings into account by some impartial principle.¹⁷
- (ii) The degree to which there is reason for me to do an action is proportional to the degree to which it promotes my good.

These are not inconsistent, because it's possible, if unlikely, that there is always a perfect correlation between the degree to which an action promotes general good and the degree to which it promotes my good. However, these statements are too weak in that they don't specify what it is that *gives* me reason to perform an action. The impartialist's thought is not just a thought about a positive correlation which happens to hold between the strength of one's reason to do something and the degree to which doing it promotes general good. His claim, rather, is that only the fact that an action will promote general good to some degree *gives* one reason to do that action—and it gives it to *that* degree. Likewise, the egoist's thought is that only the fact that an agent's action will promote that agent's good to some degree *gives* the agent reason to perform that action—and it gives it to that degree.

But put like this, the thoughts of the impartialist and the egoist are directly contradictory. So perhaps we should weaken them. We should say that the promotion of general good constitutes *a* reason for action and the promotion of one's own good also constitutes *a* reason for action. Thus:

¹⁷ The last clause leaves open the question of what impartial distributive principle should be adopted.

- (I) The fact that an action will promote to some degree the good of beings overall, taking the good of all beings into account by some impartial principle, gives anyone a reason of proportionate degree to do that action.
- (E) The fact that an action will promote the agent's own good to some degree gives that agent a reason of proportionate degree to perform that action.

Yet this must now be weaker than what Sidgwick had in mind. For if all we are saying is that (I) and (E) are both true, then it's unclear why they should come into conflict at all, and in particular, unclear why a moral government of the universe should be required to avert chaos. (I) and (E) simply specify a type of fact which is reason-giving and further specify the way in which that type of fact determines the strength of that specific type of reason.

Indeed something like (E) straightforwardly follows from (I): holding effects on other people's well-being constant, the degree to which an agent has reason to do an action varies proportionately with the degree to which that action promotes the agent's good. For the agent's own good is a constituent of general good. What was intended, however, was that (E) is somehow *independently* reason-giving. How are we to explain this intended independence? One way to do it is by giving an account of how reasons from the impartial source and reasons from the egoistic source combine. On this approach the good of the agent must have some special extra weighting in the practical reasoning of the agent.

Suppose we gave some such account. It's not clear that that would capture Sidgwick's intention. For we would now have a unified account of the principle of practical reason—a universalism with an agent-relative bias. There would be no scope for Sidgwick's worries about his dualism. Sidgwick, after all, says that 'Practical Reason' feels a 'vital need'

of proving or postulating [a] connexion of Virtue and self-interest, if it is to be made consistent with itself. For the negation of the connexion must force us to admit an ultimate and fundamental contradiction in our apparent intuitions of what is Reasonable in conduct; and from this

admission it would seem to follow that the apparently intuitive operation of the Practical Reason, manifested in these contradictory judgements, is after all illusory.¹⁸

Evidently he does not imagine that impartialism and egoism combine into a consolidated principle. In some way he sees them as *competing* perspectives on practical reason. There is a perspective on practical reason which is egoistic and another perspective which is impartial. He takes them to be equally authoritative or inescapable, a sort of permanently forced gestalt switch or bifocalism. But he also thinks that if no 'legitimately obtained conclusion or postulate as to the moral order of the world' can guarantee their coincidence, then 'the apparently intuitive operation of the Practical Reason ... is after all illusory': in 'a recognized conflict between self-interest and duty, practical reason, being divided against itself, would cease to be a motive on either side; the conflict would have to be decided between one or other of two groups of non-rational impulses'.¹⁹

3. *Supererogation*

I find this a more interesting view than universalism with an agent-relative bias. It captures the point that the egoist's and impartialist's respective claims to give an account of the basic principle of pure practical reason are *hegemonic*. Sidgwick is sensitive to that point (no doubt he is helped here by his interest in the history of moral philosophy). But on the other hand it's not easy to say how he positively sees the relation between the two principles. Perhaps he didn't really think this through. I myself cannot see how a dualism of practical reason of the kind Sidgwick envisages can be coherently formulated. Sidgwick thinks that egoism and impartialism are not a priori contradictory but can lead to contradiction in conjunction with some plausible-seeming a posteriori factual propositions. (i) and (ii) fit that picture, but as we've noted are too weak inasmuch as they don't state what *gives* one reason to act. On

¹⁸ *Methods*, 506.

¹⁹ *Ibid.*

the other hand, if (I) and (E) are so strengthened as to become hegemonic, they also become directly contradictory, irrespective of what 'moral order' there is in the world. However, I won't pursue the question²⁰ of how to formulate the dualism because I want to propose that there is no such dualism—though I also want to try to explain why there might seem to be.

Let me approach this claim by considering a case in which egoism and impartialism clearly collide. It is also a case of supererogation—and that too helps to highlight the questions we need to ask ourselves.

Fred is holed up in a defensive position with his fellow soldiers. He's fighting against an evil enemy whose victory would greatly damage the general good. Now a grenade is thrown in. Fred's got three options. His best chance of saving himself is to run away, his next best is to take evasive action by throwing himself to the ground. But he could also throw himself on to the grenade. This last action, he knows, would save the most people. For he knows that his comrades won't run away—they can see that they have an obligation to stay at the post and go on fighting, and they'll do that. But they won't throw themselves on the grenade either. So if he throws himself on it, stifling its impact, more soldiers will be saved and the chance of resisting the evil enemy will be greater. This, then, is the action that will most promote general good. Coming to this conclusion, he throws himself on the grenade.

The impartialist will endorse this reasoning. But he can of course agree that the right *moral* assessment is that Fred went beyond the call of duty. Fred and his comrades had an obligation not to run away; but they did not have an obligation to fall on the grenade. Our morality is

²⁰ For more on this question see David Phillips, 'Sidgwick, Dualism and Indeterminacy in Practical Reason', *History of Philosophy Quarterly*, 15 (1998), 57–78. Phillips suggests that Sidgwick may have held an 'indeterminacy view' according to which there is both a rational obligation to maximise one's own good and a rational obligation to maximise good impartially—except when those obligations conflict, in which case what one rationally ought to do is indeterminate. (Presumably in that case there is still a rational obligation with disjunctive content—to maximise one good or the other.) This strikes me as a philosophically unattractive position, in that it's hard to reconcile with supervenience, that is, the point that it's *in virtue* of the fact that an action maximises my good, or impartial good, that I ought to do it. Why should this supervenience lapse when my good and impartial good conflict?

not heroic; it gives some latitude to self-concern. It's also sensible, in that it would be silly to require everyone in this situation to throw themselves on the grenade. Nevertheless, according to the impartialist, if Fred had his facts right, then he was also right to think that there was more reason for him to stifle the grenade at the cost of his life than just to take evasive action. What Fred did was admirable, though not morally obligatory, and according to the impartialist it was admirable because Fred rightly saw that it *was* what there was most reason for him to do and had the courage and spirit of sacrifice to do it.

A universalist with an agent-relative bias might take a different view. If his agent-relative bias was strong enough, he could say that there was most reason for Fred to take evasive action. Neither this view nor the impartialist's view is more in line with morality, in so far as morality only says that Fred has an obligation not to run away. It certainly doesn't say that Fred has an obligation *not* to throw himself on the grenade. Only the pure egoist view, which says that Fred should run away, is inconsistent with morality.²¹ As to the bifocal view, it says that what there is most reason for Fred to do depends on whether one is focusing impartially or egoistically. Or, if we take seriously Sidgwick's view, which I quoted above—that in 'a recognized conflict between self-interest and duty, practical reason, being divided against itself, would cease to be a motive on either side; the conflict would have to be decided between one or other of two groups of non-rational impulses'—it simply ceases to provide rational guidance.²² One thing that Fred doesn't have most reason to do on this last account is what most of us would do—that is, fall to the ground.

Personally I find the impartialist account most plausible. Fred's courage and self-sacrifice are admirable. The agent-relative bias we've just considered can agree with that, but—if it gives Fred's own interest enough weight—it can't agree that Fred did the thing there was most reason for him to do. If it allocates enough strength to Fred's own interest, it will hold that there was in fact more reason, or at least as much reason, for him to take evasive action. This view can accept that he

²¹ Here I assume that one should do what one has a moral obligation to do.

²² This is in line with the indeterminacy view suggested by Phillips.

showed courage and self-sacrifice, but—inasmuch as he showed it in pursuit of what he took to be the thing there was most reason to do—it also has to hold that he acted from a mistake, whereas Fred might well have thought there was point to this courage and self-sacrifice *only* if this was *the* thing there was most reason to do. I find myself on the side of Fred. It seems to me that he was not mistaken in his motivating thought—that falling on the grenade was what there was most reason for him to do. Though he went beyond the call of duty he acted for the best, and he showed courage and self-sacrifice in doing so. No doubt I would have been among those taking evasive action; however, I don't claim either that that would have been the *best* thing for me to do or that it would have been the thing there was *most reason* for me to do.

4. *Practical reasons and their sources*

To move forward we need to consider the various sources from which reasons for acting derive. In particular, we need to consider the way in which reasons to feel give rise to reasons to act.

Consider again the case of gratitude. Out of sheer goodness of heart, someone does me an unrequested good turn. That fact certainly gives me reason to feel grateful. And because I have reason to feel grateful to him for his good turn, I have reason to express that gratitude, for example by thanking him or giving him a present or by returning the favour. Suppose on the other hand that he did me some undeserved harm. In that case I have good reason to feel resentful. And because I have reason to feel resentment I have reason to express that resentment, by recrimination, insistence on apology or even by seeking compensation.

These connections between reasons to feel and reasons to act are examples of what I have elsewhere called the 'Feeling/Disposition Principle': 'If there's reason to feel ϕ there's reason to do what feeling ϕ characteristically disposes one to do.'²³ And I think one can say that

²³ *Ethical Explorations*. By the characteristic disposition of a feeling I mean not just any action it causes but the action it functions to cause. For example, fear functions to cause flight, though it may cause one to be rooted to the spot.

the practical reasons generated by this principle do not stem from practical reason, at least when 'practical reason' is understood in a fairly common, narrow way: let us say, from 'pure practical reason'. It's not pure practical reason that tells me that if I have suffered a harm—even one that was unintended by the harmer—I have reason to demand an apology.²⁴ About this kind of normative knowledge rationalism is hardly plausible. What I need to understand, rather, is something about the hermeneutics of two emotions, resentment and regret. I need to understand that an undeserved harm can be an appropriate object of resentment on the part of the person harmed, and further, that when a person feels reasonable resentment towards me, then it is reasonable for me to feel sorry and to express that feeling by apologising to him and if necessary by reparation.

I could not know these vital normative truths (other than by testimony) if I did not myself have feelings of resentment and regret. That's what I mean when I say that these truths belong to the hermeneutics of these emotions. To know them I must know what resentment and regret are—know it, so to speak, from within. Pure practical reason can't tell me that, by definition, since by 'pure practical reason', or 'practical reason' used narrowly, we mean a faculty which enables us to recognise reasons irrespective of our capacity to feel emotions and thus to know what they are. This also means that the Feeling/Disposition Principle is not a principle of pure practical reason, since we know it by knowing its instances, and these are intelligible only to those capable of the relevant feeling.

But now consider the special case of desire. The desire for an object is a feeling whose characteristic disposition is: trying to get that object. So it falls under the Feeling/Disposition Principle in the following way:

- (D) If there is reason for you to desire *x* then there is reason for you to obtain, achieve, bring about *x*.

In this principle the notion of what there is reason for you to desire is not to be understood instrumentally, that is, as what there is reason to

²⁴ This principle is only roughly stated. If you beat me in an important competition you may in one sense at least harm me but you don't have to apologise (though you might say that you're sorry that only one of us could win, and so on).

desire because it satisfies some more basic desire. Rather, the notion invokes a distinction between reasonable and unreasonable desires—in line with the distinction between reasonable and unreasonable gratitude, resentment, envy, admiration, annoyance and so forth. In all of these cases we can often judge that a particular feeling would in a particular situation have been reasonable or unreasonable—whether or not the person in question actually felt that way. For example, it would have been quite reasonable for you to feel resentment; there was good reason for you to feel it—whether or not you felt it. A desire can be reasonable or unreasonable in the same non-instrumental way. For example, there's good reason for you to want that particular cake on the tray—you particularly enjoy chocolate cakes and that's a good one.²⁵ And just as reasonable gratitude or resentment rationalise their characteristic disposition, so reasonable desire rationalises its characteristic disposition. Notice that (D) is only a conditional, not a biconditional. I am not saying that you've got reason to obtain, achieve, or bring about *x* *only* if you've got reason to desire *x*. There may well be things you have reason to do which you don't have any non-instrumental reason to desire to do, such as throwing yourself on the grenade—indeed there often are.

So desire falls under the Feeling/Disposition Principle. But I also said it was a special case. One way in which it's special, which is particularly relevant to the present discussion, is by dint of its special connection to the notion of a person's good. It can be argued that the notion of a person, *N*'s, good is definable as that which there is reason for *N* to desire. If this definition is right, then (E) reduces to (D). Or rather it reduces to the following, strengthened, version of (D).

(D') If there is reason for *N* to desire *x* to a given degree then there is reason of proportionate degree for *N* to obtain, achieve, bring about *x*.

On this definition of a person's good the egoist's principle reduces to a special case of the Feeling/Disposition Principle. And that means that

²⁵ Your desire for that particular piece of cake need not be instrumental to a desire for—that is, it need not be a 'means to'—your own enjoyment, as has long been pointed out: for example by Bishop Butler, or J. S. Mill.

its source is not pure practical reason, but the hermeneutics of desire. Moreover, it can at least be argued that Sidgwick himself defined a person's good as what there is reason for that person to desire.²⁶ So Sidgwick too could accept the reduction of egoism to the Feeling/Disposition Principle.

But further, on this definition of a person's good we can also rewrite (I), the principle which stated that 'the fact that an action will promote to some degree the good of beings overall, taking the good of all beings into account by some impartial principle, gives anyone a reason of proportionate degree to do that action'. The good of a being, on this definition, is whatever it is that that being has reason to desire. So what (I) is saying is that if an action x , which is open to you, would promote to some degree something which some being N has reason to desire, then you have some reason to do x . How strong the reason is will depend on how strong N 's reason to desire the thing is, and how much effect, positive or negative, doing x would have on what other beings have reason to desire—taking the strength of their reasons to desire into account and computing it all by some impartial distributive principle.

Can (I) be said to be a principle of pure practical reason? That depends on how strongly one takes the notion of pure practical reason. Perhaps a principle of pure practical reason should not refer to *any* notion of what there is reason to feel—including what there is reason to *desire*—at all. Kant might insist on that, since the Categorical Imperative, derived as he envisages from the very notion of a reason for acting, makes no reference to what there is reason to feel. Correspondingly, it makes no reference to the notion of a person's good, which, we have argued, is indeed to be understood in terms of a notion of evaluative, as against practical, reason—namely the notion of what there is reason for that person to desire.

But non-Kantians are likely to feel that this is so strict a notion of pure practical reason as to be empty. So I propose a less strict notion. A principle of pure practical reason can make reference to the notion of what there is reason for an individual to desire, and hence to the notion

²⁶ See *Methods*, Bk I, ch. IX, sect 3. But see Skorupski, *Ethical Explorations*, 120, n. 23, and Skorupski, 'Desire and Will in Sidgwick and Green', *Utilitas*, 12 (2000): 316–17.

of an individual's good, but not to any other notion of what there is reason to feel. That means that on our analysis both (E) and (I) can be principles of pure practical reason. But are they?

I have already argued that (E) can be seen instead as a case of the Feeling/Disposition Principle. But (I) is different. The significant point about it is that it introduces a requirement of impartiality: it is this that takes it out of the realm of evaluative reasons into the realm of pure practical reason. There is, (I) says, reason for anyone impartially to promote the good of everyone. Famously, according to Sidgwick, this can be seen to be true 'from the point of view of the Universe'—and he has been well mocked for saying so.²⁷ But he could in my view rightly have said that it is true from the standpoint of pure practical reason—from the standpoint that is, for example, required in one's capacity as a citizen, as against the standpoint of the concretely situated individual with this or that specific endowment of affective dispositions triggered appropriately to that situation.

5. *Warrant and defeasibility*

(I), then, is a principle of pure practical reason; (E) is not. Thus there is no dualism of pure practical reason in the way in which Sidgwick envisages. But both (I) and (E), according to what has been said so far, are 'self-evident'—in that each has its own underived a priori warrant. So what happens if they come into conflict? Do we add them up or does one trump the other?

Here I can only speak for myself, certainly not for Sidgwick. It seems to me that it is the impartial principle that is finally determinative of the strength of reasons. And when its determinations can be known, they are the last court of appeal. To go back to the case of the grenade. Fred certainly has reason to desire to save himself, and everyone else in the dug-out has reason to desire to save themselves too. So

²⁷ See Bernard Williams, 'The point of view of the universe: Sidgwick and the ambitions of ethics', in his *Making Sense of Humanity* (Cambridge: Cambridge University Press, 1995): 153–71.

by (D'), that is (E), each can rightly conclude that he has reason to save himself. Is this reason, the reason from (E), somehow added to the reason from (I), so that each of the soldiers has most reason to fall to the ground, or even most reason to run away? I have already suggested that it isn't. Fred might think as follows:

We all of us have a fundamental reason to pursue our own good, and what's more it often isn't clear what action would best promote the good of all. But here it's clear. And when it's clear what best promotes the good of all then that *is* what one has most reason to do—though not necessarily the morally obligatory thing to do.

If that's what Fred thinks, then he seems to me to think truly.

(I) is not just immediately warranted—it's indefeasible. The same, I suggest, cannot be said of (E). Suppose circumstances make it reasonable for me to feel gratitude to someone and to thank them, or to feel resentment and demand an apology. However, suppose it's also clear in the circumstances that the general good will be served if I don't thank or demand an apology. Then it's plausible to hold that the reason to thank or to demand apology is overruled (though the reason to *feel* grateful or resentful is not). These reasons for action, then, are defeated by (I). Suppose on the other hand that, in the circumstances, it's not the *general* good but *my own* interest that is best served if I don't thank or demand an apology. Then it's not clear that the reason to thank or demand an apology is overruled. Here two instances of the Feeling/Disposition Principle clash. I may feel, at least to some point, that there's more reason to show gratitude than to pursue my own best interest. Or I may feel that honour requires an apology even though I see that self-interest is somewhat better served by silently accepting the insult. And I may be right. Such reasons, then, are not automatically defeated by (E).

Finally we must bring back into the picture the concepts of moral obligation and blame. It would be wrong for the soldiers to run away from the dug-out when the grenade is thrown in, but it is not wrong for them to take evading action by falling to the floor. So Fred goes beyond the call of duty in falling on the grenade. If this is what we think, then we think that any soldier who runs away is blameworthy, but that a soldier who falls to the floor is not. Obviously Fred is not blameworthy

either, so the moral obligation is disjunctive: either fall to the floor or fall on the grenade. The supererogatory element in Fred's act is that he chooses the second disjunct.

Moral wrongness, as thus illustrated, turns on blameworthiness—on what there is reason to blame. The core of blame is a certain complex of feeling—let us just call it 'the blame-feeling'. If resentment's characteristic disposition is recrimination and a call for reparation, the blame-feeling's characteristic disposition is the act of blame and the call for punishment or repentance. Just like all other feelings, the blame-feeling has its reasons which do not come from pure practical reason. But can we say that there is no connection at all between pure practical reason and the reasonableness of the blame-feeling? Compare the case of annoyance. Whether it's reasonable to be annoyed by something—a recalcitrant nail, scratching on a blackboard—is not a matter of pure practical reason at all (though of course whether it's reasonable, taking everything into account, to *act* on that annoyance may be, as we have noted). Can we say the same about the blame-feeling?

To answer this question we would have to take up some other issues in the hermeneutics of the blame-feeling. In particular we would have to ask whether it can ever be the case that a person has most reason to do something that's morally wrong. If the answer to this is, as I think it is, no, that must mean that questions about moral wrongness and questions about what impartiality requires can't be as divorced as questions about what is annoying and questions about what impartiality requires are. For, on that view, if (I) entails that some action is the one the agent has most reason to do, it will also entail that it can't be morally wrong for the agent to do it, and so entail that the agent can't be *blameworthy* in doing it. An action done because the agent reasonably thinks it is required by (I) can't be morally wrong, though such an action can certainly be 'prudentially wrong'. Hence it can't be reasonable to feel the blame-feeling towards such an action, and so considerations of pure practical reason can enter into the reasonableness of the blame-feeling. None the less it remains true that our ordinary moral intuitions are 'immediate' intuitions about blameworthiness. They are not *derived* from (I), though they are *corrigible* by (I).

6. Conclusion

To sum up, contrary to Sidgwick, there is no dualism of pure practical reason. Pure practical reason simply says that the good of any being is agent-neutrally good. On the other hand, if what we're talking about is not just pure practical reason but practical reasons in general, that is reasons for action, then there are many, and not just two. Under the Feeling/Disposition Principle an irreducible variety of reasonable feelings gives rise to an irreducible variety of underived reasons to act. Egoism has no special authority to overrule any of these practical reasons; properly understood it is simply a special case of them. On the other hand, both morality and the impartial principle have that authority; they can overrule reasons that come from the Feeling/Disposition Principle. And the impartial principle has a further authority over judgements about what there is reason to blame, which I have only alluded to briefly. What I have suggested in this case is that our judgements of blameworthiness do not *derive* from impartial reason, but that they stand open to correction by impartial reason, in a way that judgements about what is annoying, admirable, desirable for oneself and so on do not.

So Sidgwick was right to connect common-sense morality with the impartial principle; it's also understandable that he should be cagy about how it derives from that principle. However, it would be wrong to characterise this connection by saying that ordinary moral intuition cannot constitute *knowledge* until its content is derived from the impartial principle. Ordinary moral intuition can deliver immediately warranted judgements. Such moral judgements are, as a matter of general epistemological principle, defeasible, as indeed ordinary perceptual judgements are. But unless the truth of fallibilism about a domain is inconsistent with our having knowledge in that domain, there is no reason to deny that ordinary moral intuition can also deliver knowledge.