

The Methods of Ethics

Henry Sidgwick

Copyright ©2010–2015 All rights reserved. Jonathan Bennett

[Brackets] enclose editorial explanations. Small ·dots· enclose material that has been added, but can be read as though it were part of the original text. Occasional •bullets, and also indenting of passages that are not quotations, are meant as aids to grasping the structure of a sentence or a thought. Every four-point ellipsis indicates the omission of a brief passage that seems to present more difficulty than it is worth. Longer omissions are reported between brackets in normal-sized type.—The division of the work into Books, chapters, and numbered sections is Sidgwick's. —Cross-references follow this system:

‘chapter 3’ means ‘chapter 3 of *this* Book’.

‘chapter 4.2’ means ‘chapter 4, section 2, of *this* Book’.

‘II/3’ means ‘Book II, chapter 3’.

‘IV/3.4’ means ‘Book IV, chapter 3, section 4’.

An accompanying page-number refers to the page where the passage in question *starts*.—This version omits most of the 2,000+ cautions that Sidgwick includes, such as ‘I think. . .’, ‘I conceive. . .’, ‘it seems. . .’ and so on. Even with these out of the way, the work doesn't come across as bullishly dogmatic.—In this version, most notably on pages 166 and 196, the author addresses the reader (‘you’), but in the original it is always ‘the reader’ and ‘he’.—This version is based on the sixth edition of the work (1901), the last non-posthumous one. The first edition appeared in 1874, the year after Mill died.

First launched: October 2011

Contents

| | |
|--|---------------|
| BOOK I | 1 |
| Chapter 1: Introduction | 1 |
| Chapter 2: The relation of ethics to politics | 7 |
| Chapter 3: Ethical judgments | 12 |
| Chapter 4: Pleasure and desire | 19 |
| Chapter 5: Free-will | 25 |
| Chapter 6: Ethical principles and methods | 35 |
| Chapter 7: Egoism and self-love | 40 |
| Chapter 8: Intuitionism | 44 |
| Chapter 9: Good | 49 |
| BOOK II: Egoistic hedonism | 55 |
| Chapter 1: The principle and method of hedonism | 55 |
| Chapter 2: Empirical hedonism | 57 |
| Chapter 3: Empirical hedonism (continued) | 60 |
| Chapter 4: Objective hedonism and common sense | 69 |
| Chapter 5: Happiness and duty | 75 |
| Chapter 6: Deductive hedonism | 83 |
| BOOK III: Intuitionism | 93 |
| Chapter 1: Intuitionism | 93 |
| Chapter 2: Virtue and duty | 102 |
| Chapter 3: Wisdom and self-control | 109 |
| Chapter 4: Benevolence | 112 |
| Chapter 5: Justice | 125 |
| Chapter 6: Laws and promises | 141 |
| Chapter 7: Classification of duties. Veracity | 150 |
| Chapter 8: Other social duties and virtues | 154 |
| Chapter 9: Self-regarding virtues | 157 |
| Chapter 10: Courage, humility, etc. | 160 |
| Chapter 11: Review of the morality of common sense | 163 |
| Chapter 12: Motives or springs of action as subjects of moral judgment | 176 |

| | |
|---|------------|
| Chapter 13: Philosophical intuitionism | 182 |
| Chapter 14: Ultimate good | 191 |
| Book IV: Utilitarianism | 200 |
| Chapter 1: The meaning of utilitarianism | 200 |
| Chapter 2: The proof of utilitarianism | 203 |
| Chapter 3: How utilitarianism relates to the morality of common sense | 205 |
| Chapter 4: The method of utilitarianism | 223 |
| Chapter 5: The method of utilitarianism (continued) | 230 |
| Concluding chapter: The mutual relations of the three methods | 241 |

Glossary

affectation: Sometimes used here in its early-modern sense, covering every sort of pro or con attitude—desires, approvals, likings, disapprovals, dislikings, etc. Thus, the phrase ‘benevolent affections’ [page 23] isn’t pleonastic and ‘malevolent affections’ [page 154] isn’t self-contradictory.

appetite: A strong desire for some immediate end; perhaps a craving. Our narrower sense of the word is captured on page 21 by the phrase ‘appetite of hunger’.

art: Sidgwick sometimes uses ‘art’ in an older sense in which an ‘art’ is any human activity that involves techniques or rules of procedure—e.g. medicine, farming, painting.

categorical: Opposite of ‘conditional’. ‘If it won’t do anyone any harm, tell the truth’ is a conditional imperative; ‘Tell the truth!’ is a categorical imperative (see page 98; also page 4).

crucial experiment: Experiment that *settles* some question one way or the other.

Dead Sea apple: A disease-caused bulge on the bark of an oak, vaguely resembling an apple.

desert: Deservingness. The stress is on the second syllable, as in ‘dessert’ (the sweet course of a meal).

disinterested: This meant for Sidgwick what it still means in the mouths of literate people, namely ‘not *self*-interested’.

duty: Most English-language moral philosophers, Sidgwick included, speak a dialect in which ‘I have a duty to do A’ means the same as ‘I morally ought to do A’. That is not what it means in English, where ‘duty’ is tied to jobs, roles, social positions. The duties of a janitor; the duties of a landowner; ‘My Station and its Duties’ [title of a famous paper].

expedient: Advantageous, useful, helpful.

expose: In some parts of ancient Greece, unwanted babies were ‘exposed’, i.e. left out in the wilds to be killed by nature.

extra-regarding: This phrase uses ‘extra’ to mean ‘outside one’s own feelings’, and is contrasted with ‘self-regarding’. When you hang a picture, your immediate aim might be **(i)** the picture’s being on the wall or **(ii)** your enjoying seeing the picture on the wall. Of these, **(i)** is extra-regarding, **(ii)** is not.

felicific: happy-making.

generous: On page 157 Sidgwick uses this word in a sense that was dying in his day, namely that of ‘noble-minded, magnanimous, rich in positive emotions’ etc. In that passage he uses ‘liberal’ to mean what we mean by ‘generous’. Elsewhere in the work, it’s for you to decide which sense is involved.

indifference: Indifferent conduct is neither praiseworthy nor wrong; you are ‘indifferent to’ the pain of others if your thinking that a certain action would cause pain doesn’t affect your behaviour; ‘indifferent’ sensations are neither nice nor nasty.

infelicific: Not felicific.

intuition: Sidgwick uses this word in one of the two senses that it has traditionally had, in which it names the activity of (or capacity for) seeing or grasping something’s truth through a single mental act, in contrast with ‘demonstration’ which is getting there by following a proof of it. The moral position that he calls ‘intuitionism’ is the thesis that the truth or validity of some moral rules can be seen *immediately* rather than through any kind of demonstration; and thus that those rules are *basic*. See Sidgwick’s own explanation on page 44.

jural: Of or pertaining to the law.

mental: About half the occurrences of this are replacements for 'psychical'; Sidgwick evidently treats the two words as synonymous.

mutatis mutandis: A Latin phrase that is still in current use. It means '(mutatis) with changes made (mutandis) in the things that need to be changed'.

natural theology: Theology based on facts about the natural world, e.g. empirical evidence about what the 'purposes' are of parts of organisms etc.

positive: This multicoloured word is used by Sidgwick in four of its senses. **(1)** Especially in Book II, in contrast with **negative**. **(2)** In the opening paragraphs and elsewhere, in contrast with '**practical**' (with the latter including 'ethical'): a 'positive' study is one that involves no value-judgments or moral rules. **(3)** On page 71 and elsewhere, the contrast is with '**relative**': You measure a set of weights relatively if you get the facts about which is heavier than which; you measure them positively if you find out how much each weighs. Also:

positive law: On pages 8 and 15 and elsewhere this means the law of the land: a plain humanly established system of laws, in contrast with **divine law** and **moral law**. Also:

positive morality: This refers to 'the actual moral opinions generally held in a given society at a given time' (page 12). This may be a coinage of Sidgwick's (see page 101).

principles: When on page 42 Butler is quoted as speaking of 'the cool principle of self-love' he is using 'principle' in a sense that it had back in his day, in which 'principle' means 'source', 'cause', 'drive', 'energizer', or the like. (Hume's *Enquiry Concerning the Principles of Morals* is an enquiry into the *sources in human nature* of our moral thinking and feeling.)

psychogenetic: = 'having to do with the origin and development of mental states and processes'. It replaces Sidgwick's exotic 'psychogonical'.

realise: When Sidgwick speaks of 'realising' a virtue he means 'making it real', 'acting on it', 'exhibiting it in one's actions'. He explains 'self-realisation' when he uses it.

remorse: In some places these days 'remorse' means simply 'regret over something one has done' ['buyer's remorse']. In the present work it means what it once meant everywhere: '*guilty-feeling* regret over something one has done'—a sense of having acted in a morally wrong way. This is essential to an understanding of the important first paragraph of I/5.4.

requital: Pay-back: rewarding a good deed, punishing a bad one, paying a debt, etc.

sophistication: Deception by means of bad but plausible argument. So self-sophistication [page 30] is one kind of self-deception.

sympathy: From Greek meaning 'feel with': in its early modern sense, and still in Sidgwick's use, you can 'sympathise' with someone's pleasure as well as with her pain. It covers every kind of 'echo' of someone else's feelings.

tact: 'A keen faculty of perception or ability to make fine distinctions likened to the sense of touch.' (OED)

tautology: A kind of circular truth that doesn't convey any news. On page 166 Sidgwick says that a certain proposition boils down to 'Immoral acts ought not to be performed', which is a tautology because what it *means* to call an act 'immoral' is that it ought not to be performed.

Book IV: Utilitarianism

Chapter 1: The meaning of utilitarianism

1. The term 'utilitarianism' is in common use these days, and is supposed to name a doctrine or method that we're all familiar with. But it turns out to be applied to several theories that aren't logically connected with one another and don't even have the same subject-matter. So I'll do my best to make clear the doctrine that I'll call 'utilitarianism' in this Book, distinguishing it from other doctrines that could be given the same name, and indicating its relation to these.

By 'utilitarianism' I mean the ethical theory according to which

in any given circumstances the objectively right thing to do is what will produce the greatest amount of happiness on the whole

—taking into account all whose happiness is affected by the conduct. I will sometimes call this principle, and the method based on it, by the cumbersome name 'universalistic hedonism', as a reminder of what it is.

There are three confusions to be avoided. (a) Utilitarianism must be distinguished from the egoistic hedonism discussed in Book II. But the difference between

- 'each **ought to** seek his own happiness' and
- 'each **ought to** seek the happiness of all'

is so obvious that instead of dwelling on it I should explain how they ever came to be confounded or in any way included under one notion.¹ When I briefly discussed this matter in I/6, I pointed out that the confusion between these two •ethical theories was helped by confusing both with the •psychological theory that

In voluntary actions every agent **does** seek his own individual happiness or pleasure.

This has no *necessary* connection with any ethical theory; but there's some natural tendency to pass from psychological to ethical hedonism, and that transition must be primarily to the egoistic version of the latter. From the fact that everyone actually does seek his own happiness we can't conclude, as an immediate and obvious inference or even as a natural appendage, that he ought to seek the happiness of other people. (In III/13 I criticised Mill's attempt to exhibit this inference.)

(b) Utilitarianism as an ethical doctrine isn't necessarily connected with the psychological theory that the moral sentiments are derived—by 'association of ideas' or otherwise—from experiences of the non-moral pleasures and pains caused in the past, to the agent or to others, by different kinds of conduct. An intuitionist might accept this psychological theory...yet still hold that when these moral sentiments show up in our present consciousness as independent impulses they ought to have the authority they seem to claim over the more primary desires and aversions from which they have arisen. . . . In short, the so-called 'utilitarian' theory of the origin of the moral sentiments can't, unaided, prove the ethical doctrine I am calling 'utilitarianism'. I'll try in chapter 4 to show that this psychological theory has an important though subordinate place in the establishment of ethical utilitarianism.

¹ In Mill's *Utilitarianism* this confusion, though openly frowned on, is to some extent encouraged by Mill's treatment of the subject.

(c) The doctrine that universal happiness is the ultimate *standard* doesn't imply that universal benevolence is always the right or best *motive*. As I have already pointed out, the end that gives the criterion of rightness needn't always be the end that we consciously aim at; and if experience shows that general happiness will be better achieved if men frequently act from motives other than pure universal philanthropy, those other motives are preferable *on utilitarian principles*.

2. Let us now examine the utilitarian principle itself more closely. I tried in II/1 to make the notion of *greatest happiness* clear and definite; and the results of that discussion are as relevant to universalistic hedonism as to egoistic hedonism. By 'greatest happiness', then, I mean the greatest possible surplus of pleasure over pain; with equal amounts of pain and pleasure conceived as cancelling one another out for purposes of ethical calculation. Here as before it's assumed that •all pleasures included in our calculation can be compared quantitatively with one another and with pains; that •every such feeling is desirable to a certain positive or negative degree and that this degree can be to some extent known; so that •each can be roughly weighed in ideal scales against any other. This assumption is involved in the very notion of *maximum happiness*. . . .so that whatever force is given to the objections brought against this assumption in II/3 must of course tell against utilitarianism.

Who are the 'all' whose happiness is to be taken into account? Should our concern extend to all the beings capable of pleasure and pain whose feelings we can affect? or should we confine our view to *human* happiness? Bentham and Mill adopt the former view, as do (I believe) utilitarians generally; and it is obviously more in accordance with the universality of their principle. A utilitarian thinks it is his duty to aim at the good *universal*—i.e. property or quality or state—interpreted and defined as 'happiness 'or 'pleasure';

and it seems arbitrary to exclude from this project any pleasure of any sentient being.

In II/3 I pointed out the scientific •difficulties in comparing pleasures, and you may think that by broadening the scope of utilitarianism we are greatly increasing •them: if it's hard to compare the pleasures and pains of other men accurately with our own, a comparison of either with the pleasures and pains of lower animals is obviously even darker. But the difficulty isn't greater for utilitarians than for any other moralists who pay some moral attention to the pleasures and pains of lower animals. But even if we attend only to human beings, it's still not quite determinate who the morally relevant 'all' are. How far we are to consider the interests of posterity when they seem to conflict with those of now-existing human beings? The answer to this, though, seems clear: the time at which a man exists can't affect the value of his happiness from a universal point of view; so the interests of posterity must concern a utilitarian as much as those of his contemporaries—except in that the effect of his actions on the lives and even the existence of posterity must be more uncertain. ·Note 'even the existence': we can influence how many future human (or sentient) beings there will be; which raises the question of how, on utilitarian principles, this influence should be exercised. In discussing this I shall assume that for human beings generally life on the average yields a positive balance of pleasure over pain. Some thoughtful folk have denied this; but the denial conflicts with the common experience of mankind as expressed in their common patterns of action. The great majority of men, in the great majority of conditions in which human life is lived, certainly act as if death were one of the worst of evils for themselves and for those they love; and the administration

of criminal justice proceeds on a similar assumption.¹

Assuming, then, that the average happiness of human beings is a positive quantity, it seems clear that utilitarianism directs us to make the number of happy people as large as we can without lowering the average level of happiness. But if we foresee as possible that an increase in numbers will be accompanied by a decrease in average happiness, or vice versa, a point arises that •hasn't ever been explicitly discussed and •seems to have been substantially overlooked by many utilitarians—i.e. seems not to have had even a subliminal influence on their thinking. Utilitarianism prescribes as the ultimate end of action, *happiness on the whole*, not any individual's happiness except considered as a part of the whole. It follows that if the additional population enjoy on the whole positive happiness, we ought to weigh the amount of happiness gained by the extra number against the amount lost by the remainder. So that the point up to which population ought to be encouraged to increase is not that at which average happiness is the greatest possible,

as is often assumed by political economists of the school of Malthus, but

that at which the product formed by multiplying the number of persons living by the amount of average happiness reaches its maximum.

That conclusion looks rather absurd to common sense, because its show of exactness is grotesquely at odds with our awareness of the inevitable inexactness of all such calculations in actual practice. But the fact that our practical

utilitarian reasonings must be rough isn't a reason for not making them as precise as we can; and we'll be more likely to succeed if we keep clearly in mind the strict type of calculation that we would have to make if all the relevant factors could be estimated with mathematical precision.—This is a general point that is relevant to much utilitarian discussion, including the next paragraph.

It's obvious that there may be many ways of distributing the same quantum of happiness among the same number of persons; so if we are to make the utilitarian criterion of right conduct as complete as possible, we need to know which of these ways is preferable. This question is often ignored in expositions of utilitarianism. Perhaps it has seemed idle, raising a purely abstract and theoretical question that couldn't come up in everyday life; and no doubt it's true that if all the consequences of actions could be estimated with mathematical precision we probably wouldn't *ever* find the excess of pleasure over pain exactly equal for two competing lines of conduct. But just because hedonic calculations are so indefinite, it's quite likely that we should confront two sets of consequences with no difference *that we can see* between the quantities of happiness they involve. . . . And in such a case it is practically important to ask which way of distributing this quantum of happiness is the better. The utilitarian formula seems not to answer this question; it needs to be supplemented by some principle of just or right distribution of the happiness that is in question. Most utilitarians have tacitly or explicitly adopted the principle of pure equality, as given in Bentham's formula: 'Everybody to

¹ Those who hold the opposite opinion seem to assume that the appetites and desires that drive ordinary human action are in themselves painful—a view entirely contrary to my own experience and, I believe, to the common experience of mankind. See I/4.2. So far as their argument doesn't arise from that psychological error, any plausibility it has seems to come from dwelling one-sidedly on the annoyances and disappointments that certainly do occur in normal human life, and on the exceptional sufferings of small minorities of the human race, or perhaps of most men during small portions of their lives. . . .

count for one, and nobody for more than one.’ This seems to be the only principle that doesn’t need a special justification,

because—as we saw—it must be *reasonable* to treat any one man in the same way as any other if there’s no apparent *reason* for treating him differently.¹

Chapter 2: The proof of utilitarianism

When I discussed the method of egoistic hedonism in Book II, I didn’t examine any proof of its first principle; and my main concern with universalistic hedonism also is not •how its principle is to be proved to people who don’t accept it but •what logically follows from it. In fact, the principle of aiming at universal happiness is more generally felt to require some proof—some ‘considerations determining the mind to accept it’ (Mill’s phrase)—than the principle of aiming at one’s own happiness. As a matter of abstract philosophy, I don’t see why the egoistic principle should pass unchallenged any more than the universalistic one; I don’t see why the axiom of prudence shouldn’t be questioned, when it conflicts with present inclination, for a reason like the one that egoists have for rejecting the axiom of rational benevolence. If the utilitarian has to answer this:

‘Why should I sacrifice my own happiness for the greater happiness of someone else?’

then it must be all right to ask the egoist:

‘Why should I sacrifice a present pleasure for a greater one in the future? Why should I care about my own future feelings more than the feelings of others?’

Common sense finds it paradoxical to ask *why* one should seek one’s own happiness on the whole; but I don’t see how the demand can be rejected as absurd by those who belong to the extreme empirical school of psychologists, though their views are commonly supposed to be closely linked to egoistic hedonism. Grant with Hume and his followers that •the ego is merely a system of coherent phenomena, that the permanent ‘I’ is not a fact but a fiction; then why should one part of the series of feelings that constitute the ego care about •another part of the same series any more than with •any other series?

I shan’t press this question now, because I admit that common sense sees no point in giving the individual reasons for **seeking his own interest**.² Reasons for **doing his duty** according to the commonly accepted standard of duty are not seen as superfluous; utilitarian reasons are continually being given for commonly received rules of morality. Still, the fact that certain rules are commonly accepted as binding, though it does not prove them to be self-evident, removes any need to prove their authority to the common sense that accepts them; whereas a utilitarian who claims to supersede

¹ The topic is the distribution of *happiness*, not of the *means to happiness*. If some means to happiness will give more happiness to B than to A, then utilitarian principle says firmly that it ought to be given to B, whatever inequality in the distribution of the *means* of happiness this may involve.

² The relation of egoistic to universalistic hedonism is further examined in the concluding chapter of this work.

those rules by a higher principle is naturally challenged, by intuitionists as well as egoists, to prove his claim. Some utilitarians would reply by saying that it's impossible to 'prove' a first principle; and this is true if a 'proof' of P has to be a process that infers P from premises that give it its certainty; because *that* would show that these premises are the real first principles, and thus that P isn't one. But there's another difficulty—having nothing to do with the status of 'first principle'. If utilitarianism is to be proved to an intuitional moralist . . . or an egoist, the premises of the proof will have to be propositions that they accept; and that means—from the utilitarian's point of view—that the proof will have to have a conclusion that is superior in validity to the premises from which it starts. . . . How shall we deal with this dilemma? How is such a process—clearly different from ordinary proof—possible or conceivable? Yet there seems to be a general demand for it. What is needed, perhaps, is a line of argument which •allows *some* validity to the **maxims already accepted** but also •shows that they aren't absolutely valid and need to be controlled and completed by **some more comprehensive principle**. [Sidgwick could have written, more specifically, ' . . . a line of argument which allows some validity to **egoism and intuitionism** but also shows that they aren't absolutely valid and need to be controlled and completed by **utilitarianism**'.]

I gave such a line of argument, addressed to egoism, in III/13. Note, though, that whether this argument works depends on how the egoistic first principle is formulated. If the egoist strictly confines himself to saying that he ought to take his own happiness or pleasure as his ultimate end, he leaves no opening for any reasoning to lead him to universalistic hedonism as a first principle;¹ the difference

between •his own happiness and •other people's happiness is all-important *for him*, and there's no way to argue him out of this. All the utilitarian can do in that case is to try to reconcile the two principles. . . .by pointing out to the egoist the pleasures (pains) that he can expect to have if he observes (violates) the utilitarian rules. This may incline him to seek the greatest happiness of the greatest number as •a means to his own happiness—but not of course as his ultimate end. So it's nothing like a *proof* of utilitarianism. But if the egoist says or implies that his happiness or pleasure is good, not only for him but from the point of view of the Universe—e.g. by saying that nature designed him to seek his own happiness—it is relevant to tell him that his happiness can't be a more important part of good, taken universally, than the equal happiness of anyone else. In this way he may be brought to accept universal happiness as absolutely and unqualifiedly good or desirable—as an end that a reasonable agent ought to pursue.

This is the reasoning I used in III/13 when exhibiting the principle of rational benevolence as one of the few intuitions that stand the test of rigorous criticism.² Notice though that when this argument is addressed to the intuitionist, it shows only that the utilitarian first principle is *one* moral axiom; it doesn't prove that it is sole or supreme. The premises the intuitionist starts with include other formulae that he thinks are also independent and self-evident. So utilitarianism has to be related in two ways, one **negative** and the other **positive**, to these formulae. **Negatively** the utilitarian must try to show to the intuitionist that the principles of truth, justice, etc. have only a dependent and subordinate validity; arguing either

¹ He may be led to it in other ways—by appeals to his sympathies, or to his moral or quasi-moral sentiments.

² The argument in III/13 leads to the first principle of utilitarianism only if it's admitted that happiness is the only thing ultimately and intrinsically good or desirable. I tried in III/14 to get common sense to make this admission.

- that really common sense affirms the principle only as a general rule admitting of exceptions and qualifications, as in the case of truth, and that a further principle is needed to systematise these exceptions and qualifications; or
- that the fundamental notion is vague and needs to be made more precise, as in the case of justice; that the different rules are liable to conflict with each other and a higher principle is needed to settle these conflicts; and that. . . .there are many signs of the vagueness and ambiguity of the common moral notions to which the intuitionist appeals.

I have given this part of the argument in Book III. Now I must supplement this line of reasoning by developing the **positive** relation between utilitarianism and the morality of common sense. I have to do this by showing how utilitarianism

supports the general validity of current moral judgments, making good the defects that reflection finds in the intuitive recognition of their stringency; and at the same time provides a method for binding the unconnected and sometimes conflicting principles of common moral reasoning into a complete and harmonious system. If systematic reflection on the morality of common sense can in this way show the utilitarian principle to be what common sense naturally appeals to for the further development that this same reflection shows to be necessary, that will give utilitarianism as complete a proof as it is capable of. Can this project succeed? To answer that we need to study something that is also important in its own right, namely the exact relation of utilitarianism to the commonly received rules of morality. I shall address this at some length in the next chapter.

Chapter 3: How utilitarianism relates to the morality of common sense

1. I pointed out in I/6 [page 39] that the two sides of utilitarianism's relation to the morality of common sense have been prominent at two different periods in the history of English ethical thought. Since Bentham we have mainly seen the negative or aggressive aspect of utilitarianism. But when Cumberland in replying to Hobbes said that the received moral rules generally tend to promote the 'common good

of all rational beings', his aim was simply conservative; it never occurred to him to consider whether these rules are imperfect, or whether common moral opinions disagree with the conclusions of rational benevolence.¹ So in Shaftesbury's system the 'moral sense' is supposed to be always pleased with the 'balance' of the affections that tends to the good or happiness of the whole, and displeased with the opposite.

¹ Cumberland doesn't adopt a hedonistic interpretation of good. But I follow Hallam in regarding him as the founder of English utilitarianism; because it seems that 'good' came gradually and half-unconsciously to have the definitely hedonistic meaning that it has implicitly in Shaftesbury's system and explicitly in Hume's.

Hume treats this topic in detail and with a more definite assertion that the moral likings (or aversions) aroused in us by different qualities of character and conduct all come from the perception of utility (or the reverse).¹ Notably, the most penetrating critic among Hume's contemporaries, Adam Smith, fully accepts that •rightness objectively coincides with •utility; though he maintains against Hume that 'our view of this utility is neither the first nor the principal source of our approval'. Of Hume's theory that

'the only qualities of the mind that are approved of as virtuous are ones that are useful or agreeable to the person himself or to others, and the only ones that are disapproved of are those with a contrary tendency'

Smith agrees, and remarks that 'Nature seems to have adjusted •our sentiments of approval and disapproval to •the convenience of the individual and of the society so happily that after the strictest examination it will be found, I believe, that Hume's thesis is universally true'.

And no-one can read Hume's *Inquiry into the Principles of Morals* without being convinced of at least this much: If we made a list of the qualities of character and conduct that directly or indirectly produce pleasure to ourselves or to others, it would include all that are commonly known as virtues. Whatever the origin of our notion of moral goodness, there's no doubt that •utility is a general characteristic of the dispositions we describe as good, and that •to that extent it's true that the morality of common sense is at least unconsciously utilitarian. You might object:

This coincidence •between goodness and utility• is merely general and •qualitative; it breaks down when we attempt to fill in the details with the

•quantitative precision that Bentham introduced into the discussion.

The assertion that virtue always produces happiness is indeed very different from the assertion that the right action is always the one that which will produce the greatest possible happiness on the whole. But remember that utilitarianism isn't concerned to prove that the intuitional and utilitarian methods absolutely **coincide** in their results. (And if it *could* prove that much, this success would be almost fatal to its practical claims because it would mean that it doesn't make the slightest difference whether one adopts the utilitarian principle.) What utilitarians are called on to show is a natural **transition** from the morality of common sense to utilitarianism, somewhat like the transition in (say) bridge-building from •trained instinct and empirical rules to •the technical method that provides a scientific basis for the activity; so that utilitarianism can be regarded as the scientifically complete and systematically thought-out form of the regulation of conduct that has through the whole course of human history tended in the same general direction. It doesn't need to prove that existing moral rules do *more* for general happiness than any others; but only to point out in each case some clearly felicitous [see Glossary] tendency that they have.

But Hume's dissertation exhibits, along the way, more than a simple and general harmony between common moral sentiments regarding actions and the actions' foreseen pleasurable and painful consequences. The utilitarian argument can't be fairly judged without *fully* taking into account the cumulative force that it gets from the complex nature of its coincidence with common sense.

It can be shown, I think, that •the utilitarian estimate of

¹ Hume's sense of 'utility' is narrower than Bentham's, and more in accordance with ordinary language. He distinguishes the 'useful' from the 'immediately agreeable'; and holds that there are some elements of personal merit that we approve because they are 'immediately agreeable' to the person who has them or to others. But it's more convenient here to use the word in the wider sense that has been current since Bentham.

consequences relates to the current moral rules by supporting not only the general outlines but also their commonly accepted limitations and qualifications; that •it explains anomalies in the morality of common sense, anomalies that must from any other point of view seem unsatisfactory to the reflective intellect; that •faced with the difficulties and perplexities that arise when people try to remedy the imprecision of the current rules, it solves these in a manner that generally agrees with common sense's vague instincts and is naturally appealed to in ordinary moral discussions of these difficulties; that •it not only supports the common sense view of the relative importance of different duties, but is naturally called in as arbiter when rules come into conflict; that •when one rule is interpreted a little differently by different persons, each naturally supports his view by urging its utility, however strongly he maintains that the rule is self-evident and known *a priori*; that •when there's marked diversity of moral opinion on any point, in one country at one time, we commonly find obvious and impressive utilitarian reasons on both sides; and finally that •most of the remarkable discrepancies found among the moral codes of different ages and countries are strikingly correlated with differences in the effects of actions on happiness, or in men's foresight of such effects or their care about them. Hume makes most of these points, in a somewhat casual and fragmentary way; and many of them were incidentally illustrated in my examination of common sense morality in Book III. But because of the importance of this matter, I should exhibit in systematic detail the cumulative argument that I have summed up in this paragraph, even at the risk of repeating some of the results previously given.

2. Here's an objection that is frequently urged against utilitarianism:

'If the true ground of the moral goodness or badness of actions lies in their utility or the reverse, how can we explain common sense's broad distinction between the moral part of our nature and the rest of it? Why is the excellence of virtue so strongly felt to be different in kind. . . .from people's physical beauties and aptitudes and their intellectual gifts and talents?

I answer this by saying—as I did in III/2—the only qualities that are strictly *virtuous* are ones we think can be realised at least to some extent by voluntary effort, so that the conspicuous obstacle to virtuous action is the lack of an adequate motive. So we expect that judgments of moral goodness, passed by the agent or by others, will supply a fresh motive on the side of virtue and thus have an immediate effect in causing actions to be at least externally virtuous; and the habitual awareness of this will account for any difference between •moral sentiments and •the pleasure we get from contemplating non-voluntary utilities. To this, however, it is replied the tendencies to strictly voluntary actions include many that aren't commonly regarded as virtuous and yet are more useful than many virtues.

'The selfish instinct that leads men to •accumulate does more good for the world than the generous instinct that leads men to •give. . . . A modest, diffident, and retiring nature, distrustful of its own abilities and humbly shrinking from conflict, benefits the world less than does the self-assertion of an audacious and arrogant nature that is impelled to every struggle and develops every capacity. Gratitude has done much to soften and sweeten the intercourse of life, but the opposing feeling of revenge was for centuries the one bulwark against social anarchy and is still one of

the main restraints to crime. On the great stage of public life, especially in great convulsions where passions are fiercely roused, the man who confers most benefit on the world isn't the delicately scrupulous and sincerely impartial man, or the single-minded religious fanatic who can't deceive or delay. It is rather the astute statesman, earnest about his goals but unscrupulous about his means, equally free from the tangles of conscience and the blindness of zeal, who governs—because he partly yields to the passions and the prejudices of his time. But. . . it has scarcely yet been maintained that the delicate conscience that in these cases detracts from utility constitutes vice!' (W.E.H. Lecky, *History of European Morals*, chapter 1.

This is forceful but not, it seems to me, very difficult to answer. Bear in mind, though, that the present argument doesn't aim to prove that utilitarian inferences coincide exactly with the intuitions of common sense, but only aims to show that those intuitions are primitively and imperfectly utilitarian.

Firstly: Let us distinguish the recognition of goodness in dispositions from the recognition of rightness in conduct. An action that a utilitarian must condemn as likely to do more harm than good may come from a disposition that will on the whole produce more good than harm. This is often the case with scrupulously conscientious acts. However true it is that unenlightened conscientiousness has driven men to fanatical cruelty, mistaken asceticism, and other non-felicific conduct, I don't think any intuitionist would deny that care in conforming to accepted moral rules has an over-all tendency to promote happiness. Note that when we see a generally felicific disposition having unhappy results in a particular case, we often apply to it some term of condemnation—e.g. we say that the person has been

'over-scrupulous' or 'fanatical'. But that is consistent with our regarding that same disposition as a good element of character. **Secondly**, although a utilitarian holds that only what's useful is praiseworthy, he doesn't have to maintain that *how* worthy of praise something is depends strictly on *how* useful it is. To repeat a point I made earlier: from a utilitarian point of view what we must mean by saying that a quality 'deserves praise' is that it's expedient [see Glossary] to praise it with a view to its future production. In distributing utilitarian praise of human qualities our chief concern is with the usefulness not of the •quality but of the •praise; and it's obviously not expedient to encourage by praise qualities that are likely to be found in excess. Self-love and resentment, for example, are necessary to society, but it's quite in harmony with utilitarianism that common sense doesn't recognise them as virtues. . . . But when self-love conflicts with impulses that are on the whole pernicious, it is praised as 'prudence'; and when a man seems clearly deficient in resentment, he is criticised for being 'tame'; but it's natural that the occasional utility of malevolent impulses is somewhat overlooked, given how obviously productive of pain they are. Something like this holds also for humility and diffidence. As I showed in III/10, it's a careless mistake for common sense to praise the tendency to underrate one's own powers; most people when they think about it agree that it can't be good to be in error about this or anything else. But the desires for superiority and esteem are so strong in most men that arrogance and self-assertion are much commoner than the opposite defects, and they are also faults that are specially disagreeable to others. That is why humility gives us a pleasant surprise, and common sense is easily led to overlook the more latent and remote bad consequences of undue self-distrust.

The morality of common sense seems to be perplexed about how •moral excellence relates to •moral effort, but this is cleared up when we adopt a utilitarian point of view. On one hand, it's easy to see how some acts are likely to be more felicific when performed without effort and from motives other than regard for duty; while on the other hand, someone who in performing such acts achieves a triumph of duty over strong seductive inclinations exhibits a character that we recognise as felicific in a more general way. . . . There's also a simple and obvious utilitarian solution of the problem of whether we should influence someone to do something •right• that he thinks is wrong. A utilitarian would weigh •the felicific consequences of the particular right act against the •infelicific [see Glossary] results likely to come from the moral deterioration of the person if other motives lead him to act against his conscientious convictions. . . . And I think that that's the calculation that the common sense of mankind would also conduct, in a vague and semiconscious way.

But if we are to estimate precisely how far utilitarianism agrees with common sense, it seems best to examine judgments of right and wrong in conduct under the headings represented by our common notions of virtues and duties. Let me first remind you that these common notions •aren't rivals to utilitarianism•: when adequately precise definitions of them turn out to involve the notion of 'good' or of 'right' supposed to be already determinate, they have no basis for opposing a utilitarian interpretation of 'good' or 'right'. For example, *wisdom* is not commonly conceived as the faculty of choosing the right means to the end of universal happiness; rather, as we saw in III/3, the common notion of wisdom involves an uncritical synthesis of the different ends and principles that are sorted out and separately examined in this work. But if its meaning isn't distinctly utilitarian, it certainly isn't anything else either; so that the definition

leaves it open to us to give the notions *good* and *right* a utilitarian import.

3. Let us start with the virtues and duties discussed in III/4 under the heading of **benevolence**. As regards the general conception of the duty of benevolence, I don't think there's any significant divergence between the intuitional and utilitarian systems. Benevolence might be more commonly defined as a disposition to promote the *good* of one's fellow-creatures, which involves not merely their *happiness* as utilitarians understand that but also their *moral good or virtue*; but if we can show that the other virtues are all generally conducive to the happiness of the agent himself or of others, it will follow that benevolence. . . . aims directly or indirectly at the utilitarian end. (Notice that I am not here assuming the conclusions of III/14 in their full breadth.)

And the comprehensive range that utilitarians give to benevolence—aiming at the greatest happiness of all sentient beings—seem not to be really opposed to common sense. Some intuitional moralists do restrict the scope of the direct duty of benevolence to human beings, and regard our duties to lower animals as merely indirect and derived from the duty of self-culture, but it's they who appear paradoxical, rather than their utilitarian opponents. In saying that each agent is to think of all other happiness as being as important as his own, utilitarianism seems to •go beyond the standard that is commonly set under the heading of benevolence, but it can't be said to •conflict with common sense on this point. The **practical** application of this theoretical impartiality of utilitarianism is limited by at least two important considerations. **(a)** Generally speaking, each man can provide for his own happiness better than he can for that of others—because he knows more about his own desires and needs, and has greater opportunities to gratify them. **(b)** The stimulus of self-interest is what most easily and thoroughly draws out

the active energies of most men; if it were removed, general happiness would be lessened •by a serious loss of the means of happiness that are obtained through labour; and even to some extent •by the lessening of the labour itself. For these two reasons it wouldn't promote universal happiness if each man concerned himself with the happiness of others as much as with his own. Whereas if I consider the duty of benevolence **abstractly and ideally**, even common sense morality seems to bid me to love my neighbour as myself.

Here is a plausible objection to utilitarianism:

'Under the notions of generosity, self-sacrifice, etc., common sense *praises*. . . a suppression of egoism beyond what utilitarianism approves. We may admire as virtuous a man who gives up his own happiness for someone else's sake, even if the happiness that he confers is clearly less than that which he resigns so that there's a lessening of over-all happiness.'

I have three replies to this. **(a)** I don't think we *do* entirely approve of such conduct when the sacrifice/benefit ratio is obviously large. **(b)** A spectator often can't judge whether over-all happiness is lost, because he cannot tell •how far the benefactor is compensated by sympathetic and moral pleasure, or •what remoter felicific consequences may come from the sacrifice's moral effects on the agent and on others. **(c)** Even if there *is* a loss in the particular case, our admiration for the self-sacrifice may be justifiable on utilitarian grounds: such conduct shows a disposition that will generally tend to promote happiness, and it is may be this disposition that we admire rather than the particular act.

Some critics have said that the rigid impartiality of the utilitarian formula ignores the special claims and duties arising from each man's special relations to a few out of the whole number of human beings; and hence that although

utilitarianism and common sense may agree that all right action is beneficent in being conducive to the happiness of someone or other, they diverge on the radical question of the *distribution* of beneficence.

It seems that on this point even fair-minded opponents have misunderstood the utilitarian position. They have attacked Bentham's well-known formula *Every man to count for one, nobody for more than one* on the ground that the general happiness will be best attained by inequality in the distribution of each one's services. But if it's clear that it will be best attained in this way, utilitarianism will prescribe this way of aiming at it! Bentham's dictum doesn't lay down a rule of conduct; it merely aims says that when we are computing how much happiness a given state of the world involves, we should give equal weight to any two equally happy people. And it's pretty obvious why it is generally conducive to the general happiness that each individual should distribute his beneficence in the channels marked out by commonly recognised ties and claims. There are two reasons for this.

(i) In the chief relations discussed in III/4—the domestic, and those constituted by blood-relatedness, friendship, previous kindnesses, and special needs—the services that common sense prescribes as duties are commonly •prompted by natural affection, and they also tend to •develop and sustain such affection. The existence of benevolent affections among human beings is itself an important means to the utilitarian end, because (as Shaftesbury and his followers forcibly urged) the most intense and highly valued of our pleasures come from such affections—the emotion itself is pleasurable and adds pleasure to the activities that it prompts. . . . [Sidgwick goes on at some length about the thesis that 'spontaneous beneficence' is risky because it tends to make the beneficiary passive. He says that this

bad effect is much less likely if] the alms are bestowed with unaffected sympathy and kindness, and in such a way as to elicit a genuine response of gratitude. . . .

That is why the utilitarian will approve of the cultivation of affection and the performance of affectionate services. It may be said that what we ought to approve is not so much •affection for special individuals but rather •a feeling more universal in its scope—charity, or philanthropy. It's true that special affections will occasionally conflict with the principle of promoting the general happiness; so utilitarianism must prescribe a culture of the feelings that will counteract this tendency. But it seems that most persons are capable of strong affections towards only a few people in certain close relationships, especially the domestic, so that if these were suppressed we would feel towards our fellow-creatures generally a 'watery kindness' (Aristotle's phrase) that would be a feeble counterpoise to self-love! So the specialised affections that our society normally produces provide the best means of developing in most people a more extended benevolence, as far as they are capable of that. Besides, hardly anyone has the power or the knowledge to do much good to many people; and that is in itself a reason why it's desirable that our chief benevolent impulses should be correspondingly limited.

(ii) The second reason it is conducive to the general happiness that special claims to services should be commonly recognised as attaching to special relations doesn't concern affection as such. . . . We saw in III/4.1 that where there are these relations common sense regards the affection itself as a duty but still prescribes the performance of the services even if the affection is absent. The services that we are commonly prompted to by the domestic affections or by gratitude or pity

are indeed an integral part of the system of mutual aid by which the normal life and happiness of society is maintained, as an indispensable supplement to the still more essential services that are definitely prescribed by law or explicitly promised in contracts. Political economists have explained to us that the means of happiness are immensely increased by that complex system of co-operation that has been gradually organised among civilised men; and although it is thought that in such a system it is generally best for an individual to exchange •such services as he is willing to provide for •whatever return he can get for them by free contract, there are many large exceptions to this general principle. The most important ones concern children. The well-being of mankind requires that in each generation children should be

- produced in adequate numbers, neither too many nor too few;
- adequately nourished and protected during the period of infancy; and
- carefully trained in good intellectual, moral, and physical habits;

and it is commonly believed that the best—or even the only known—means of achieving these ends is provided by the existing institution of the family, resting as it does on a basis of legal and moral rules combined. Law fixes a minimum of mutual services and draws the broad outlines of behaviour for the different members of the family, imposing¹ on the parents lifelong union and complete mutual fidelity and the duty of providing the necessities of life for their children up to a certain age; in return for which it gives them the control of their children for the same period, and sometimes lays on the children the burden of supporting their parents when they are aged and destitute; so that when morality enjoins a

¹ The law of modern states doesn't outright *enforce* this; but it refuses to recognise domestic partnership contracts of any other kind, and the social effect is substantially the same.

completer harmony of interests and a fuller interchange of kindnesses, it is merely filling in the outlines drawn by law. When we tried to formulate the domestic duties recognised by common sense we found in most cases a vague margin with regard to which there is •no general agreement and •continual disputes. And now I point out that the latent utilitarianism of common moral opinion shows up most clearly in this margin; for when there's a dispute about the precise mutual duties of husbands and wives, for example, or of parents and children, the disputants usually support their views by predicting the effects on human happiness of the general establishment of their proposed rule; this seems to be the standard that is applied by common consent.

Natural sympathy moves us to recognise the claim for help to those who are in special need; a *moral* basis for such claims can obviously be provided by utilitarianism; indeed the meeting of them seems so important to society's well-being that in most modern civilised communities the law has something to say about them. I noted that the main utilitarian reason why it's not right for every rich man to distribute his superfluous wealth among the poor is that it's best for the over-all happiness of everyone that adults generally (except married women) should expect that each will have to find ways of meeting his own wants. But if I discover that because of a sudden and unforeseeable calamity someone's resources are clearly inadequate to protect him from pain or serious discomfort, the case is altered: my •theoretical obligation to consider his happiness as much as my own becomes •practical, and I'm obliged to make as much effort to relieve him as won't involve greater loss of happiness to myself or others. If the calamity could have been foreseen and averted by proper care, my duty becomes more doubtful; for then by relieving him I risk encouraging others to be improvident. In such a case a utilitarian has to weigh this

indirect evil against the direct good of removing pain and distress; and it is now increasingly widely recognised that the question of providing for the destitute—whether by law or by private charity—has to be treated as a utilitarian problem of which these are the elements.

Cases where it is conducive to general happiness that one man x should render unbought services to another man y are not only ones where y is poor. Whatever a man's economic status, he may find himself unable to •ward off some evil, or to •bring about some worthy end, without help of a kind that he can't purchase in the labour-market; and it may be help that won't have a bad effect on him (because this is an exceptional emergency) and won't be burdensome to the giver. Some legal theorists have thought that where the service is great and the burden slight, it might be proper for the service to be required by law—so that if I could save a man from drowning by merely holding out a hand, I would be legally punishable if didn't do that. Be that as it may, the *moral* rule condemning the refusal of aid in such emergencies is obviously conducive to the general happiness.

The need for unbought services isn't confined to emergencies. There are other services for which there is normally no market-price—e.g. advice and assistance in the intimate perplexities of life, which one is willing to receive only from genuine friends. Rendering such a service brings emotional pleasure to the benefactor, and also contributes to general happiness in other ways. That is why we see friendship as an important means to the utilitarian end. Yet we feel that the charm of friendship is lost if the flow of emotion is not spontaneous and unforced. The combination of these two views seems to be exactly represented by the sympathy that is not quite admiration with which common sense regards all close and strong affections; and the regret that is not quite disapproval with which it contemplates their decay.

Whenever it is conducive to general happiness that unbought services should be rendered, gratitude—meaning a settled disposition to repay the benefit when and how one can—is demanded by utilitarianism no less than by common sense. [Sidgwick goes on at length about this, focusing on an ‘apparent puzzle’:] •Benefits conferred without expectation of reward have a peculiar excellence. . . ., but •it would be difficult to treat as a friend someone from whom gratitude was not expected. . . . This is one of the cases where an apparent ethical contradiction turns out to be a mere matter of psychological complexity. Most of our actions are done from several motives, so that this can happen:

A man has a disinterested desire to help his friend, a desire that would prevail even if he had no hope of requital; but this generous impulse is sustained by a vague trust that requital won’t be withheld.

The apparent puzzle provides another illustration of the latent utilitarianism of common sense. On one hand: utilitarianism tells us to render services whenever it is conducive to general happiness to do so, which will often be the case quite apart from any gain to oneself that would result from their requital. On the other hand: the actual selfishness of average men tells us that such services wouldn’t be adequately rendered if requital were not expected, and so it is conducive to general happiness that men should recognise a moral obligation to repay them.

I have discussed only the most conspicuous of the duties of affection; but it is probably obvious that similar reasonings would hold also for the others.

The commonly received view of special claims and duties arising out of special relations, though *prima facie* •opposed to the impartial universality of the utilitarian principle, is

really •supported by a well-considered application of that principle. Three distinct lines of argument support this claim. **(a)** Morality is here in a manner protecting the normal channels and courses of natural benevolent affections; and the development of such affections is of the highest importance to human happiness, as a direct source of pleasure and as a preparation for a broader altruism. **(b)** The mere fact that such affections are normal causes •an expectation of the services that are their natural expression; and the disappointment of •these is inevitably painful. **(c)** We can show in each case strong utilitarian reasons why services should usually be rendered to the persons commonly recognised as having such claims rather than to others.

The difficulties that we found in the way of determining by the intuitional method the limits and the relative importance of these duties are reduced in the utilitarian system to difficulties of hedonistic comparison.¹ For each of the preceding arguments has shown us different kinds of pleasures gained and pains averted by the fulfilment of the claims in question. . . . These different pleasures and pains combine differently, and with almost infinite variation as circumstances vary, into utilitarian reasons for each of the claims in question. None of these reasons is absolute and conclusive, but each has its own weight while being liable to be outweighed by others.

4. I pass to consider another group of duties, often contrasted with those of benevolence, under the comprehensive notion of **justice**.

‘That justice is useful to society’, says Hume, ‘it would be a superfluous undertaking to prove’; what he tries to show at some length is ‘that public utility is the sole origin of justice’; and the same question of *origin* is Mill’s chief topic in chapter

¹ In chapters 4 and 5 I’ll discuss further the method of dealing with these difficulties in their utilitarian form.

5 of *Utilitarianism*. My topic here, however, is not so much

- the growth of the sentiment of justice from experiences of utility, as
- the utilitarian basis of the fully grown notion.

But if my previous account of it is correct, the justice that is commonly demanded and inculcated is more complex than these writers have recognised. What Hume means by 'justice' is what I would call 'order', taking that in its widest sense, as referring to

the observance of the actual system of legal and customary rules that bind the members of a society into an organic whole, checking injurious impulses, distributing the different objects of men's clashing desires, and demanding such positive services as are commonly regarded as *owed*, whether through contract or by custom.

There have always been plausible empirical arguments for the revolutionary thesis that 'laws are imposed in the interest of rulers', but Hume is still right: the general conduciveness to social happiness of the habit of order or law-observance is too obvious to need proof. Indeed, order is so important to a community that even if a particular law is clearly injurious it is usually expedient to obey it, apart from any penalty the individual might suffer from breaking it. We saw, however, that common sense sometimes tells us to refuse obedience to bad laws, because 'we ought to obey God rather than men' (though there seems to be no clear intuition about the kind or degree of badness that justifies resistance); and it also allows us in special emergencies to violate rules that are generally good, because 'Necessity has no law' and 'The well-being of the people is the highest law'.

These and similar common opinions suggest that the limits on the duty of law-observance are to be determined by utilitarian considerations. And the utilitarian view gets rid of

the difficulties we encountered in trying to define intuitively the truly legitimate source of legislative authority (see III/6.2-3); while it also justifies to some extent each of the current views about the intrinsic legitimacy of governments.

For obedience: Utilitarianism finds the moral basis of any established political order primarily in its effects rather than its causes, so that obedience will usually be due to any *de facto* government that isn't governing very badly.

Possibly against obedience: if laws that originate in manner M are likely to be •better or •more readily obeyed, it is a utilitarian duty to aim at introducing M; and thus in a certain stage of social development it may be right that a 'representative system' should be demanded by the people or even (in extreme cases) introduced by force.

For obedience again: It can be expedient to maintain an ancient form of legislation because men readily obey such; and loyalty to a dispossessed government ·such as that of Charles I· may be on the whole expedient, even at the cost of some temporary suffering and disorder, so that ambitious men ·such as Oliver Cromwell· don't find usurpation too easy.

Here again utilitarianism supports the different reasons commonly put forward as absolute, and also brings them theoretically to a common measure so that we have a principle of decision between conflicting political arguments in particular cases.

This obedience to law, at least when it affects the interests of other individuals, is what we often mean by 'justice'. But it seems (see III/5) that the notion of *justice* analyses out into several elements combined in a somewhat complex manner. Let us investigate now what latent utilities are represented by each of these elements.

A constant part of the notion of justice, which is there even when *just* isn't distinguished from *legal*, is impartiality, i.e. the negation of arbitrary inequality. As we saw in III/13.3, this impartiality. . . . is merely a special case of the wider maxim that it can't be right to treat differently two persons who are similar in all significant respects. And we saw that utilitarianism admits this maxim no less than other systems of ethics. But this negative criterion doesn't provide a complete determination of what laws or actions are just; so we still have to ask: What are the inequalities, in laws and in the distribution of pleasures and pains outside the sphere of law, that *aren't* arbitrary and unreasonable? and to what general principles can they be reduced?

We can explain on utilitarian principles why apparently arbitrary inequality in a certain part of individual conduct isn't regarded as unjust or even (in some cases) as open to any criticism (see footnote on page 127). Freedom of action is an important source of happiness to those who have it, and a socially useful stimulus to their energies; so it's obviously expedient that a man's free choice in distribution of wealth or kind services should usually not be restrained by fear of •legal penalties or •social disapproval; and therefore, when clearly recognised claims are satisfied, it is expedient that an individual's mere preferences should be regarded by others as legitimate grounds for distributing his property or services unequally. . . .

Let us now consider the general principles that seem to be at work in common sense's recognition of 'just claims'. The grounds for many such claims fall into the category of 'normal expectations'; and obligations in such cases vary greatly in strictness depending on whether the expectations are based on definite undertakings, or on some vague mutual understanding, or are merely such as an average man would form in those circumstances. In these latter cases

common sense seems to be somewhat perplexed, but for the utilitarian the difficulty disappears. He will hold that any disappointment of expectations is *prima facie* bad, but *how* bad in a given case depends on how confident that expectation was: the more sure he was, the greater shock he will get from the disappointment—I mean a shock to his reliance on the conduct of his fellow-men generally—and so the worse the disappointment will be. And it will be much worse still if the expectation is generally recognised as normal and reasonable, because then there is a shock not only for him but for anyone else who knows about this disappointment. It's so important to people to be able to rely on each other's actions that in ordinary cases scarcely any advantage can counterbalance the harm done by violating absolutely definite undertakings. Still, we found in III/6 that several exceptions and qualifications to the rule of good faith [= 'promise-keeping'] were fairly clearly recognised by common sense; and most of these have a utilitarian basis that it's not hard to see. I'll now sketch four of these. **(a)** The superficial view that the obligation of a promise depends on the assertion of the promiser—and not, as utilitarians hold, on the expectations produced in the promisee—can't fairly be attributed to common sense; which doesn't condemn a breach of promise so strongly when no-one has acted in reliance on it—e.g. when a man breaks a pledge of total abstinence from alcohol. So perhaps there's little if any conflict here between common sense and utilitarianism. The next three points involve an even clearer harmony between the two. **(b)** Utilitarian reasons for keeping a promise are weakened by a relevant change of circumstances (see III/6.8), because in that case the disappointed expectations are not the ones that the promise created. **(c)** It would obviously be bad for the community if men could rely on the keeping of promises procured by fraud or unlawful force, because that

would encourage the use of fraud or force for this purpose. . . .
(d) In that same section we saw that common sense •is disposed to admit that a promise isn't binding if keeping it would harm the person to whom it was made, and •isn't sure that it is binding if keeping it would *greatly* harm the promiser; and both of these qualifications are in harmony with utilitarianism. Similarly for the other qualifications and exceptions; they all turn out to be clearly supported by utilitarianism. . . .

It is undeniably a good thing for normal expectations to be satisfied even when they aren't based on a definite contract—it is clearly conducive to the settled and well-adjusted activity on which social happiness greatly depends. Utilitarianism is useful here: it spares us the difficulties that beset the common view of just conduct as something absolutely precise and definite. In this vaguer region we can't sharply demarcate valid claims from valid invalid ones; 'injustice' shades gradually off into mere 'hardship'. In practice common sense is forced to adopt the utilitarian view that the disappointment of natural expectations is an evil that must sometimes be put up with for the sake of a greater good, though it's hard to reconcile this with the theoretical **absoluteness** of justice in the intuitional view of morality.

When we examine the general conceptions of 'ideal justice' (as I have called it) that we find involved in current criticisms of the existing order of society, we become even more aware of the advantage of utilitarianism's view that the obligation to be just is **relative** [i.e. varies in strength depending on the circumstances].

The loose notions of ordinary men seem to fluctuate between two competing views of an ideally just social order—two extremes that I have called respectively 'individualistic' and 'socialistic'. According to the former of these, an

ideal system of law should aim at freedom, i.e. perfect non-interference among all the members of the community, as an absolute end. There are obvious and striking utilitarian reasons for leaving each rational adult free to seek happiness in his own way:

- each is best qualified to provide for his own interests, because even when he doesn't know best what they are and how to achieve them, he is at any rate most keenly concerned for them;
- the consciousness of freedom and the responsibility that goes with it increases the average effective activity of men; and
- the discomfort of constraint is directly an evil and *prima facie* to be avoided.

Still, we saw in III/5.4 that the attempt to devise a consistent code of laws taking maximum freedom (instead of happiness) as an absolute end leads to startling paradoxes and insoluble puzzles; and in fact no society—not even the freest ones—has in practice interpreted 'freedom' in that absolute way; every society's thoughts about freedom have been more or less consciously determined by considerations of expediency. So it's fair to say that common sense in adopting the individualistic ideal in politics has always subordinated it to and limited it by the utilitarian first principle.

But it seems that what we commonly want under the name of 'ideal justice' is not so much •freedom as •the distribution of good and evil according to desert. Indeed this is often said to be what freedom is *for*, the idea being that if we protect men from mutual interference each will reap the good and bad consequences of his own conduct, and so be happy or unhappy in proportion to his deserts. In particular, it has been widely held that with a free exchange of wealth and services each individual will obtain from society whatever money etc. his services are really worth. But we saw that the

price an individual gets for wealth or services that he sells in a system of perfect free trade may, for several reasons, *not* be proportioned to the social utility of what he is selling; and thoughtful common sense seems—under the influence of utilitarian considerations—to accept this disproportion as to some extent legitimate. Here as elsewhere, utilitarianism corrects the thoughtless utterances of moral sentiments.

For example, if a moral man is asked ‘How far is it right in bargaining to take advantage of the other party’s ignorance?’ his first impulse would probably be to answer ‘Not right at all’. But reflection would show him that this is too sweeping; that in a case like this—

x in negotiating with a stranger y takes advantage of y’s ignorance of facts that x knows and that y could have known if he had used as much diligence and foresight as x did

—common sense doesn’t blame x for this. Why not? Because we have a more or less conscious sense that restricting the free pursuit and exercise of economic knowledge is likely to lead to loss to the wealth of the community. And for somewhat similar reasons of general expediency, if the question be raised whether it is fair for a class of persons to gain by the unfavourable economic situation of any class with which they deal, common sense at least hesitates to censure such gains at any rate when such unfavourable situation is due ‘to the gradual action of general causes, for the existence of which the persons who gain are not specially responsible. [Much of this paragraph has, as Sidgwick reports, been quoted from a longer discussion in his *Principles of Political Economy* III/9.]

The general principle of ‘requiting good desert’, so far as common sense really accepts it as practically applicable to the relations of men in society, is broadly in harmony with utilitarianism, because it’s obvious that rewarding men

for felicitous conduct is favourable to general happiness. The utilitarian scale of rewards will take into account not only the value of the services performed but also the difficulty of getting men to perform them; but this element seems also to be taken into account (perhaps unconsciously) by common sense; for we don’t usually recognise merit in right actions of kinds that men are naturally inclined to perform too much rather than too little (see section 2 above and in III/2.1). Another example: the intuitional principle that ill-desert lies in wrong intention conflicts with the utilitarian view of punishment as purely preventive, but in the actual administration of criminal justice, common sense is forced into reluctant practical agreement with utilitarianism. After a civil war it demands the execution of purely patriotic rebels; and after a railway accident it demands severe punishment for unintentional neglects which would have been regarded as trivial if it weren’t for their consequences.

But in any distribution of pleasures and privileges, or of pains and burdens, where considerations of desert don’t come in (i.e. if the good or evil to be distributed has no relation to any conduct on the part of the persons concerned) or where it is impossible in practice to take such considerations into account, common sense falls back on *simple equality* as the principle of just distribution. And we’ve seen that in such a case the utilitarian will reasonably accept equality as the only method of distribution that isn’t arbitrary. In fact, this way of distributing the means of happiness is likely to produce more happiness on the whole. Why? Partly because men have a disinterested dislike of unreason, but more because they dislike being in any way inferior to others, especially when the inferiority seems unreasonable. This feeling is so strong that it often prevails in spite of obvious claims of desert. Perhaps it is sometimes expedient that it should so prevail.

Utilitarianism also provides a common standard to which the different elements in the notion of *justice* can be reduced. Such a standard is urgently needed because these different elements are continually at risk of conflicting with each other. The political issue between conservatives and reformers often involves such a conflict. If my analysis of the common notion of justice is sound, the attempt to extract from it a clear answer to this—

‘Ought we to do some violence to expectations arising naturally out of the existing social order in order to bring about a more ideally just distribution of the means of happiness?’

is certain to fail because the conflict is, so to speak permanently latent in the very core of common sense. The utilitarian will merely use this notion of *justice* as a guide to different kinds of utilities; and when these are incompatible he’ll balance one set of advantages against the other and decide according to how the scales tilt.

5. The duty of **truth-speaking** is sometimes taken as a striking instance of a moral rule that doesn’t rest on a utilitarian basis. But if you look carefully at how the common opinion of mankind actually preaches this duty you’ll see that this is not so; the general utility of truth-speaking is too obvious to need proof, and whenever this utility seems to be absent or outweighed by bad consequences, common sense at least hesitates to enforce the rule. For example, it is *prima facie* harmful to the community for a criminal to be helped in his pursuits by being able to rely on the assertions of others. So deception is *prima facie* legitimate as a protection against crime; but when we consider the bad effects that a single lie might have (by contributing to a habit of lying, and by setting a bad example), we see that the utilitarian case for the lie is doubtful; and that’s just what common sense thinks. Another example: it is generally in a man’s

interests to know the truth, but sometimes that is harmful to him—e.g. when an invalid hears bad news—and in these cases common sense is disposed to suspend the rule. An other point: we found it difficult to say exactly what veracity consists in—

- truth in the spoken words?
- truth in the inferences that the speaker thinks will be drawn from his words?
- truth in both?

Perfect candour would require it in both; but in the various circumstances where this seems inexpedient, we often find common sense at least half-willing to dispense with one or other part of the double obligation. A respectable school of thinkers maintain that a religious truth may properly be communicated by means of an historical fiction; and the common rules of politeness often require us to suppress truths and suggest falsehoods, thereby acknowledging that perfect frankness isn’t a good fit with our existing social relations. [Sidgwick adds that in most such cases common sense is a little unsure about what to allow, and says that the same is true of utilitarianism.]

The •different views people have about the legitimacy of **malevolent impulses**—making it hard for us to formulate a consistent common-sense doctrine about this—exactly correspond to •different forecasts of the consequences of gratifying such impulses. *Prima facie* the desire to injure some particular person is inconsistent with a deliberate purpose of benefiting as much as possible people in general; and so we find *superficial common sense* sweepingly condemning all such desires. But a study of the actual facts of society shows that resentment plays an important part in the socially valuable repression of injuries; so the thoughtful moralist shrinks from ruling it out entirely. But personal ill-will is obviously a very *dangerous* means to

general happiness: its immediate goal is the exact opposite of happiness; and though the achievement of this may sometimes be the lesser of two evils, it's still the case that if this impulse is encouraged it is likely to cause the infliction of pain beyond the limits of just punishment, and to harm the angry person's character. This inclines the moralist to prescribe that indignation be directed always against •actions and not against •persons. Now, it might seem that anger thus restricted would be the state of mind most conducive to general happiness *if* it would be effective in repressing injuries. But *could* the average person abide by this restriction, •always directing his anger at the action rather than the agent? And even if he could and did, *would* this redirected anger be effective enough on its own? It's not obvious that Yes is the right answer to either question, which is why common sense hesitates to condemn personal ill-will against wrongdoers even if it includes a desire for the enjoyment of seeing them suffer.

As for **temperance**, **self-control**, and the so-called **self-regarding virtues** generally, it's easy to show that they are 'useful' to the person who has them; and if common sense isn't quite clear about what the goal is of regulation and control of appetites and passions that moralists have so much preached and admired, there is at least no obstacle to holding that the goal is happiness. Even in the ascetic extreme of self-control that has sometimes led to the rejection of sensual pleasures as radically bad we can trace an unconscious utilitarianism. The ascetic condemnation has always aimed mainly at the pleasures that are especially liable •lead to excesses dangerous to health or •to interfere with the development of other faculties and susceptibilities that are important sources of happiness.

6. The regulation of the sexual appetite, prescribed under the notion of **purity** or **chastity**, seems to be an exception

to what I have been saying; because under this heading we find notably vigorous and severe condemnation of acts of which the immediate effect is pleasure not obviously outweighed by subsequent pain. But a more careful look at this 'exception' transforms it into an important •contribution to my argument, showing a specially complex and delicate correspondence between moral sentiments and social utilities.

(i) The special intensity and delicacy of the moral sentiments that govern sexual relations are thoroughly justified by the importance to society of the end to which they are obviously a means—the preservation of the permanent unions that are thought to be necessary for the proper rearing and training of children. That is why the first rule for this part of life is the one that •directly secures conjugal fidelity; and there are obvious utilitarian grounds for protecting marriage •indirectly by condemning all extra-marital sexual affairs: if the moral censure of such affairs were removed,

- men's motives for taking on the restraints and burdens of marriage would be seriously weakened;
- young people of both sexes would form habits of feeling and conduct that would tend to unfit them for marriage; and
- if extra-marital intercourse were fertile, it would lead to imperfect care of the succeeding generation. . . .

(ii) Common-sense morality has always had two views about the simple offence of unchastity—one for men and another for women—and this difference is anomalous. The offence is commonly more deliberate in the man, who has the additional guilt of soliciting and persuading the woman; and in the woman it is much more often prompted by some motive that we rank higher than mere lust; so that by the ordinary standards of intuitional morality unchastity ought to be more severely condemned in the man. Yet the common-sense attitude is the exact opposite of this, and we

look for a justification for this inversion. Only utilitarianism can provide it. It depends on the fact that society's interests are more closely tied to there being a high standard of *female* chastity. [If a wife plays around, Sidgwick explains, the husband is unsure if he is the father of their children, and that 'strikes at the root of family life'; whereas a husband's being sexually unfaithful, though it lessens the family's well-being, doesn't threaten its very existence.]

Still, the common moral sense of Christian countries these days pretty clearly and explicitly condemns •unchastity in men; though we recognise the existence of a laxer code—the so-called morality of 'the world'—which treats •it as very mildly wrong or not wrong at all. But the difference between the two codes gives a kind of support to my argument because it corresponds to a difference between more and less intelligent ways of viewing the consequences of maintaining certain moral sanctions. 'Men of the world' think that •men can't in practice be restrained from sexual indulgence, at least at the time of life when the passions are strongest; and hence that •it is expedient to tolerate illicit sexual intercourse of a kind and degree that isn't directly dangerous to the well-being of families. Some of these men, in bolder antagonism to common sense, maintain that the existence of a limited amount of such intercourse (with a special class of women, carefully separated from the rest of society as they actually are) is scarcely a real evil, and may even be a positive gain in respect of general happiness; for continence may be somewhat dangerous to health, and certainly involves a fairly intense loss of pleasure. 'The 'man of the world' defends the existence of such a class of women as follows: The maintenance of a satisfactory population-size in an old society doesn't require that *all* the women in each generation should become mothers of families; and if some of the surplus make it their profession

to enter into casual and temporary sexual relations with men, there's no *need* for their lives to be less happy than those of other women in the less favoured classes of society.

This is superficially plausible, but it ignores the social benefits of the present practice of subjecting unchaste women to severe penalties of social contempt and exclusion, resting on moral disapproval. •It keeps the class of courtesans [here = 'prostitutes'] sufficiently separate from the rest of female society to prevent the contagion of unchastity from spreading. •And it keeps the illicit intercourse of the sexes within limits so as not to interfere significantly with the due development of the race. This consideration is enough to make a utilitarian support the established rule against this kind of conduct, and therefore to condemn violations of the rule as over-all infelicitous even though they may seem to be infelicitous only because of the moral censure attached to them. The 'man of the world' is also ignoring the vast importance to the human race of maintaining the higher type of sexual relations that isn't generally possible except where a high value is set on chastity in both sexes. From this point of view the virtue of *purity* can be seen as providing a needed shelter under which the intense and elevated affection between the sexes, which is most conducive to individual happiness and to the well-being of the family, may grow and flourish.

Now we can explain something that must have perplexed many thoughtful people contemplating the common-sense regulation of conduct under the heading of *purity*, namely the fact that

- the sentiment that supports these rules is very intense, so that the subjective difference between right and wrong in this department is especially strongly marked; and yet
- it is found to be impossible to say clearly just *what* conduct is being condemned under this notion.

The impulse to be restrained is so powerful and so receptive to stimulants of all kinds that the sentiment of purity has to be very keen and vivid if it is to do its protective job; and the aversion to impurity must extend far beyond the acts that primarily need to be prohibited, and include in its scope everything—in dress, language, social customs, etc.—that might excite lustful ideas. And the line between right and wrong in such matters doesn't need to be drawn with theoretical precision; it's enough for practical purposes if the main central portion of the region of duty is brightly lit while the margin is left in shadows. Also, the detailed regulations that society needs to maintain depend so much on habit and association of ideas that they vary greatly from age to age and from country to country.

7. I have •given several illustrations of how utilitarianism is normally introduced as a method for deciding between conflicting claims where common sense leaves their relative importance obscure—e.g. between the different duties of the affections, and the different principles that turn out to be involved in our common conception of justice—and I have •shown how, when there's a dispute about the precise scope and definition of any current moral rule, it is usually thought that the dispute should be decided by the effects of different interpretations of the rule on general happiness or social well-being. Actually these two lines of thought practically coalesce, because it's generally a conflict between maxims that impresses men with the need for precise definitions. You may say:

'The "consequences" that are commonly referred to in such cases are effects on •social well-being rather than on "general happiness" as this is understood by utilitarians; so the two notions ought not to be identified.'

I grant this; but I tried in III/14 to show that when common

sense is dealing with the aspects of ultimate good or well-being that seem at first sight to be *furthest* from anything like pleasure or happiness, it nevertheless comes at these in an unconsciously utilitarian manner. And I now add that this hypothesis of 'unconscious utilitarianism' explains the fact that different classes of human beings differ in •how they rank various virtues and in •how much importance they attach to some individual virtues. For such differences ordinarily correspond to differences of view about the *utilitarian* importance of the virtues under different circumstances. I have already noted the greater stress laid on chastity in women than in men; courage is more valued in men because they are more called on to cope energetically with sudden dangers. For similar reasons, a soldier is expected to show a higher degree of courage than (say) a priest. One more example: we value candour and scrupulous sincerity in most persons, but we scarcely look for them in a diplomat who has to conceal secrets, nor do we expect a tradesman to tell his customers about the defects in his good.

Differences in the moral codes of different ages and countries correspond, at least to a large extent, to differences either in the actual effects of actions on happiness, or in the extent to which such effects are generally foreseen or regarded as important by the people concerned. I have already noted several instances of this; and the general fact, which has been emphasised by utilitarian writers, is also admitted and even emphasised by their opponents. Thus Dugald Stewart in his *Active and Moral Powers* II/3 stresses the extent to which the moral judgments of mankind have been modified by 'the diversity in their physical circumstances', the 'unequal degrees of civilisation that they have attained', and 'their unequal measures of knowledge or of capacity'. He points out that theft is regarded as a very minor offence in the South Sea Islanders, because little or no labour is

required there to support life; that lending money for interest is commonly looked down on in societies where commerce is imperfectly developed, because in such communities the 'usurer' is commonly in the odious position of wringing a gain out of the hard necessities of his fellows; and that where the legal arrangements for punishing crime are imperfect, private murder is either justified or treated lightly. More examples could be given; but few people who have studied this subject will deny that variations in •the moral code from age to age are to some extent correlated with variations in •the real or perceived effects on general happiness of actions dealt with by the code. And in proportion as the apprehension of consequences becomes more comprehensive and exact, we may trace not only change in the moral code handed down from age to age, but progress in the direction of a closer approximation to a perfectly enlightened utilitarianism. Only we must distinctly notice another important factor in the progress, which Stewart has not mentioned: the extension, namely, of the capacity for sympathy in an average member of the community. The imperfection of earlier moral codes is at least as much due to defectiveness of sympathy as of intelligence; often, no doubt, the ruder man did not perceive the effects of his conduct on others; but often, again, he perceived them more or less, but felt little or no concern about them. Thus it happens that changes in the conscience of a community often correspond to changes in the extent and degree of the sensitiveness of an average member of it to the feelings of others. Of this the moral development historically worked out under the influence of Christianity affords familiar illustrations.¹

I'm not maintaining that the development of current morality is *perfectly* correlated with the changes in the sympathy with which people have viewed the consequences of conduct. On the contrary, the history of morality shows us many signs of what seem from the utilitarian point of view to be partial aberrations of the moral sense. But even here we can often discover a germ of unconscious utilitarianism; the aberration is often only •an exaggeration of an obviously useful sentiment, or •the extension of it by analogy with cases to which it doesn't properly apply, or •what's left of a sentiment that was once useful but now isn't. [Note that Sidgwick regularly uses 'useful' as a pointer towards utilitarianism = 'useful-ism'.] Please notice that I have been careful *not* to say that the perception of the rightness of any kind of conduct has always—or even usually—been derived by conscious inference from a perception of consequent advantages. This hypothesis is naturally suggested by the survey I have conducted, but the evidence of history doesn't give it much support: as we track back in the history of ethical thought, we find that the further back we go the *less* aware common moral consciousness is of the utilitarian basis of the morality that was current at the time in question. For example, Aristotle saw that the virtue of *courage* as recognised by the common sense of ancient Greece was restricted to dangers in war: and we can *now* explain this limitation in terms of the utilitarian importance of this kind of courage at a time when the individual's happiness was tied to the welfare of his state more tightly than it is now, and when the very existence of the state was more frequently imperilled by hostile invasions; but this explanation lies well outside the range of Aristotle's own

¹ The *current* morality of the Graeco-Roman civilised world is the outcome of the extension and intensification of sympathy due to Christianity. Changes brought about in this way include: •the severe condemnation and eventual suppression of the practice of exposing [see Glossary] infants; •effective abhorrence of the barbarism of gladiatorial combats; •immediate moral mitigation of slavery and a strong encouragement of emancipation; •a great extension of the charitable provision made for the sick and poor.

thoughts. The origin of our moral notions and sentiments lies hidden in the obscure regions of *hypothetical history*, where conjecture has free scope; but when our backward look approaches the borders of this realm we *don't* find it easier to trace a conscious connection in men's minds between accepted moral rules and foreseen effects on general happiness. Early man's admiration of beauties or excellences of character seems to have been as direct and unreflective as his admiration of any other beauty; and the strictness of law and custom in primitive times seems to rest on the evils that divine displeasure will supernaturally inflict on their violators, rather than on even a rough and vague forecast of the natural bad consequences of non-observance. So that

the most reasonable claim utilitarianism can make about its relation to common-sense morality is not that it is where mankind began but rather what mankind has always been tending towards—it is the adult form of morality, not the new-born.

[In all printings of this work since 1901 a passage lifted from Book I is inserted at this point. Sidgwick had removed it from I/2 (his posthumous editor reports), intending to incorporate it in Book IV, but died before completing the revision of the work. It's hard to see *where* in this Book the passage would fit. In the present version it returns to the chapter that contained it in the editions that Sidgwick did supervise throughout, starting on page 11.]

Chapter 4: The method of utilitarianism

1. If I have sufficiently established the view I have been maintaining about the general utilitarian basis of the morality of common sense, we can now address the question: What method of determining right conduct will the acceptance of utilitarianism lead to in practice? The most obvious method is that of empirical hedonism (see II/3), according to which we have in each case to adopt the conduct that seems likely to lead to the greatest happiness on the whole.

In Book II, however, we found much perplexity and uncertainty in this method, even in the restricted application of it that we were considering there—namely, even when the agent has only to consider his own happiness. Even when someone is occupied only in forecasting his own pleasures, it seems hard or impossible for him to avoid quite big errors,

whether in •accurately comparing the pleasantness of his own remembered past feelings, or •going by the experience of others, or •arguing from the past to the future. And the difficulties increase when we have to consider the effects of our actions on all the sentient beings who may be affected by them. But I couldn't, in Book II, find any satisfactory substitute for this method of empirical comparison. It didn't seem reasonable to take refuge in the uncriticised beliefs of men in general regarding the sources of happiness; indeed, it seemed impossible to extract any clear and definite consensus from the confused and varying utterances of common sense on this subject. [Sidgwick now mentions a couple of difficulties encountered in the discussion of egoistic hedonism in II/3–4, and then:] But when we consider the accepted principles of

morality in relation to the happiness not of the individual but of human (or sentient) beings generally, it's clear from chapter 3 that the problem of harmonising hedonism with intuitionism starts to look quite different. Indeed, from the materials that I have presented in that chapter it's only a short and easy step to the conclusion that **the morality of common sense is in fact a body of utilitarian doctrine**; that the 'rules of morality for the multitude' are to be regarded as 'positive beliefs of mankind regarding the effects of actions on their happiness',¹ so that the apparent 'first principles of common sense can be accepted as the 'middle axioms' of utilitarian method; with utilitarian considerations being explicitly mentioned only in settling issues on which the verdict of common sense is obscure and conflicting. On this view the traditional controversy between the advocates of virtue and the advocates of happiness would seem to be at length harmoniously settled.

The arguments for this view that I have presented receive support from the hypothesis, now widely accepted, that moral sentiments are derived by a complex and gradual process from experiences of pleasure and pain. Briefly stated, the hypothesis seems to be this [the numbering is Sidgwick's]:

- (1) In each person's experience the pain or alarm caused to •him by actions of his own or by others tends by association to cause him to dislike such actions, and a weaker version of the same thing happens in relation to pain or alarm caused to •others to whom he has some special connection of blood or community of interest, or some special tie of sympathy.
- (2) Experience also tends to give him sentiments that restrain him from actions that are painful or alarming to others—through his dread of their resentment and

its consequences, especially dread of his chief's anger and. . . of the anger of supernatural beings.

- (3) These feelings of dread combine with a sympathetic aversion to the pain of other men generally; this is comparatively feeble at first, but tends to strengthen as morality develops. In the same way •experiences of pleasure and gratitude, and •desire for the goodwill of others and its consequences, tend to make him like actions that cause pleasure to himself or to others. So similar aversions and likings are produced in most members of any society (because they are generally alike in their natures and circumstances), and they tend to become *more* similar through communication and imitation; and individual divergences are repressed by each person's desire to retain the goodwill of others. This leads to the gradual development of common likings for conduct that gives pleasure to the community or to some part of it, and common dislikes for conduct causing pain and alarm. These (dis)likings are passed on down the generations, partly perhaps by physical inheritance but mainly by •parents instructing children and •imitation of adults by the young. In this way their origin becomes obscured, and they finally appear as what are called *moral sentiments*.

When I reflect on my own moral consciousness—my own faculty of moral judgment and reasoning—what I find doesn't square with this theory. I don't find any apparent intuitions that stand the test of rigorous examination except ones that are too abstract and general to have a recognisable relation to particular experiences—I mean the abstract principles of prudence, justice, and rational benevolence as defined in III/13. But I see no reason to doubt that the theory is

¹ See Mill's *Utilitarianism* chapter 2. But Mill says that the 'rules of morality for the multitude' are to be accepted by the philosopher only provisionally, until he has something better.

partly true about the historical origin of particular moral sentiments, habits, and commonly accepted rules; and thus add something to the arguments of chapter 3 that tend to exhibit the morality of common sense as unconsciously or 'instinctively' utilitarian.

But it is one thing to hold that current morality expresses—partly consciously but mostly unconsciously—the results of human experience regarding the effects of actions; it is a very different thing to accept this morality *en bloc*, so far as it is clear and definite, as the best guide we can have to the attainment of maximum general happiness. This simple reconciliation of intuitional and utilitarian methods may be very attractive, but it isn't warranted by the evidence. **Firstly:** It emphasises the effect of sympathy with the feelings that *result from* actions while neglecting sympathy with the impulses that *lead to* actions. Adam Smith (in Book I of his *Theory of the Moral Sentiments*) assigns to this operation of sympathy—the echo (so to speak) of each agent's passion in the breast of spectators—the first place in determining our approval and disapproval of actions; and he treats sympathy with the effect of conduct on others as a merely secondary factor, correcting and qualifying the former.¹ Without going as far as this, I'm sure there are many cases where the resulting moral consciousness looks like a balance or compromise between the two kinds of sympathy; and that compromise

can easily be far from the rule that utilitarianism would prescribe. [Sidgwick's next page is heavy going and needlessly difficult. He defends his point about the compromise on the grounds that there's no reason to expect input feelings to correlate strictly with resultant feelings, and the latter are what utilitarianism cares about. He then silently drops the 'compromise' idea and gives reasons—partly repeated from chapter 3—why people aren't very good at estimating what pleasures or pains will result from their actions. •People are limited in their degree of sympathy with the feelings of others; and Sidgwick presumably holds that you may underestimate the pain you are causing by your shortage of sympathy with it. •People are also limited in how much they know; they aren't cognitively equipped to make good judgments about causes and effects, including action-causes and feeling-effect. Sidgwick continues:] •Where the habit of obedience to authority has become strong, moral sentiments may be perverted by a desire to win the favour or avert the anger of superiors. •False religions also have influence; the sensibilities of religious teachers have influenced their followers' moral codes on matters where these sensibilities were not normal and representative, but exceptional and idiosyncratic.²

Secondly: We must suppose that these deflecting influences have been limited and counteracted by the struggle for

¹ The operation of sympathy is strikingly illustrated in the penal codes of primitive communities, both by the mildness of the punishments inflicted for homicide, and by the startling differences in penalties for the same crime depending on whether the criminal was taken in the act or not. Sir Henry Maine writes: 'It is curious to observe how completely the men of primitive times were persuaded that the injuries of the injured person were the proper measure of the vengeance he was entitled to exact, and how strictly they fixed the scale of punishment according to the probable rise and fall of his passions.' (*Ancient Law*, chapter 10) And even in more civilised societies there's a common feeling of uncertainty about the propriety of inflicting punishment for crimes committed long ago, which seems traceable to the same source.

² This influence is limited, because no authority can permanently impose on men regulations that are flagrantly infelicitous. Even the most original religious teachers have produced their effect mainly by giving new force and vividness to sentiments that men already had and recognised as authoritative in the society on which they acted. Still, human history might have been very different if, for example, Mohammed had been fond of wine, and indifferent to women.

existence in past ages, because any moral habit or sentiment that was unfavourable to the survival of the social organism would be a disadvantage in the struggle for existence, and would therefore tend to perish with the community that adhered to it. But we have no reason to suppose that this force would keep positive morality [see Glossary] always in line with a utilitarian ideal. Imperfect morality would be only one disadvantage among many, and seldom the most important one, especially in the earlier stages of social and moral development when the struggle for existence was most vigorous. Also, a morality could be •perfectly preservative of a human community while also being •imperfectly felicific, and thus in need of considerable improvement from a utilitarian point of view. And however completely adapted the moral instincts of a community are at some time to its conditions of existence, the adaptation could be ruined by some change in the community's circumstances. Apart from any visible changes in external circumstances, there might be some law of human development such that the most completely organised experience of human happiness in the past would give us little guidance in making it a maximum in the future. . . . When we turn from these abstract considerations to history, and examine the actual morality of other ages and countries, we find that morality has been an obviously imperfect instrument for producing general happiness; so there's surely a strong presumption that our own moral code has similar imperfections that habit and familiarity have hidden from us.

Thirdly: The divergences that we find when we compare the moralities of different ages and countries exist side by side in the morality of any one society at any given time. I pointed out earlier that when divergent opinions are held by a minority so large that we cannot fairly regard the majority dogma as the plain utterance of 'common sense',

there has to be an appeal to some higher principle, and very often it's utilitarianism. But a smaller minority than this, especially if it's composed of persons with •enlightenment and •special familiarity with the effects of the conduct judged, can reasonably inspire us with a distrust of common sense; just as with more technical activities we prefer the judgment of a few trained experts to the instincts of the vulgar. Thinking about these divergent codes and their relation to the different circumstances in which men live suggests that common-sense morality is really right only for ordinary men in ordinary circumstances—though it may be expedient that these plain folk should regard it as absolutely and universally prescribed, since any other view of it may dangerously weaken its hold over their minds. To the extent that *this* is how things stand, we must use the utilitarian method to discover how far persons in special circumstances require a morality better suited to them than common sense is willing to concede; and also how far men of special physical or mental constitution should be exempted from ordinary rules, as has sometimes been claimed for men of genius, or intensely emotional men, or men gifted with unusual prudence and self-control.

Fourthly: [Sidgwick says that when people are aware of a conflict between their moral views and their beliefs about what utilitarianism would say, it *may* be that] this discrepancy would disappear after a deeper and completer examination of the consequences of actions. . . . But how far would they get with this? We can't answer *a priori*, so this is really a further argument for a comprehensive and systematic application of a purely utilitarian method.

I conclude that we can't take the moral rules of common sense as expressing the consensus of competent judges up to now regarding the kind of conduct that is likely to produce the greatest happiness on the whole. It seems to be the

unavoidable duty of systematic utilitarianism to review these rules thoroughly so as to discover how far the causes I have described (and perhaps others) have actually operated to make common sense diverge from a perfectly utilitarian code of morality.

2. But that way of stating the problem assumes that the second item in the comparison—a perfectly utilitarian code of morality—can be defined and developed well enough for us to formulate with adequate precision a utilitarian moral code for human beings. This seems to have been commonly assumed by the utilitarians. But when we really try to construct such a system, we encounter serious difficulties. Setting aside the uncertainties involved in any comparison of pleasures, let us suppose that the amount of human happiness that will result from any plan of behaviour can be ascertained exactly enough for practical purposes in advance of the plan's being put into operation. It still has to be asked: What is the nature of the human beings for whom we are to make this plan? *Humanity* isn't something that exhibits the same properties always and everywhere. Whether we consider intellect, feelings, physical condition, or circumstances, we find men to be so different at different times and places that it seems absurd to lay down a set of ideal utilitarian rules for mankind generally. You may say:

These are differences only in the details. There's still enough uniformity in the nature and circumstances of human life, always and everywhere, to make possible an outline scheme of ideal behaviour for mankind at large.

I reply that *details* are precisely what we are now principally concerned with. The previous discussion has shown well enough that the conduct approved by common sense has a general resemblance to conduct that utilitarianism would prescribe; but now we want to discover more •exactly how

far the resemblance extends, and with how •precisely the current moral rules are suited to the actual needs and conditions of human life.

Let us then narrow the scope of investigation and try only to discover the rules appropriate to men as we know them, in our own age and country. We're immediately met with a dilemma. The men we know have a more or less definite moral code.

- If we think about them as having this code, we can't at the same time think of them as beings for whom a code is yet to be constructed from the ground up; but on the other hand
- if we take an actual man—e.g. an average Englishman—and set aside his morality, what remains is an entity that is so purely hypothetical that it's not clear what practical good can be done by constructing a system of moral rules for the community of such beings.

To amplify the second limb of this dilemma: Could we assume that the scientific deduction of such a system would ensure its general acceptance? Could we reasonably expect to convert all mankind—or even all educated and reflective mankind—to utilitarian principles, so that all preachers and teachers would aim at universal happiness as unquestioningly as physicians aim at the health of the individual body? Could we be sure that men's moral habits and sentiments would adjust themselves to these changed rules at once and without any waste of force? If the answer is Yes to each question, then perhaps we could construct the utilitarian code while leaving existing morality out of account. But I can't think that we are justified in making these suppositions; I think we have to take •the moral habits, impulses, and tastes of men as material given us to work on, just as much as •the rest of their nature; and because that material

only partly results from reasoning in the past, it can be only partly modified by any reasoning that we now apply to it. It seems therefore clear that we can't get a practically serviceable moral code by constructing an ideal morality for men conceived to be as men actually are except for setting aside their actual morality.

You may say:

'No doubt such an ideal utilitarian morality can be introduced only gradually, and perhaps after all imperfectly; but still it will be useful to construct it as a pattern to which we can approximate.'

But **(i)** it may not be really possible to approximate to it: any existing moral rule, though not the ideally best even for existing men in existing conditions, may still be the best that men can be induced to obey; so that proposing any other would be futile—or even harmful because it might tend to impair old moral habits without effectively replacing them by new ones. **(ii)** The attempt gradually to approximate to a morality constructed on the supposition that the non-moral part of existing human nature remains unchanged, may lead us astray; because

- the state of men's knowledge and intellectual faculties,
- the range of their sympathies,
- the direction and strength of their prevailing impulses,
- and

•their relations to the external world and to each other are continually being altered, and such alteration is to some extent under our control and may be highly felicitous; and any significant change in important elements and conditions of human life may require corresponding changes in established moral rules and sentiments, so that the human being whose life is thus modified may achieve the greatest possible

happiness. In short, the construction of a utilitarian code, regarded as an ideal towards which we are to progress, is met by a second dilemma. •If the topic is long-term planning, human nature and the conditions of human life can't usefully be assumed to be constant. •If we are attending only to the short term, men's actual moral habits and sentiments won't be significantly changeable within our time frame.

In the concluding chapters of his *Data of Ethics* Herbert Spencer maintains that the problems of practical ethics can be solved by •constructing the final perfect form of society, towards which the process of human history is tending; and •working out the rules of behaviour that ought to be, and will be, followed by the members of this perfect society. I don't accept this. For one thing, granting that we can conceive as possible a human community that is perfect by utilitarian standards, and granting also Spencer's definition of this perfection—namely that the voluntary actions of all the members cause 'pleasure unmixed with pain' to all who are affected by them¹—it still wouldn't be remotely possible to forecast the natures and relations of the members of such a community with enough clearness and certainty to be able to define even in outline their moral code. Also, even if we *could* construct Spencer's ideal morality scientifically, the construction wouldn't help us much in solving the practical problems of actual humanity. A society in which—to take just one example—there is no such thing as punishment must be one whose essential structure is so unlike ours that it would be idle to attempt any close imitation of its rules of behaviour. It might be best for us to conform approximately to *some* of these rules; but we could know this only by examining each particular rule in detail; we would have no reason to think that it would be best for us to conform

¹ Not that this definition is acceptable to a utilitarian. A society might be 'perfect' according to this definition, and yet not contain the greatest possible happiness; for there might be an even higher level of happiness which would involve a slight alloy of pain.

to all of them as far as possible. If this ideal society is going to be realised eventually, that will have to happen by evolution through a considerable period of time; so it's likely enough that •the best way of progressing towards it will be something other than the seemingly most direct way, and that •we'll reach it more easily if we begin by moving away from it. Whether and to what extent this is so can't be known except by carefully examining the effects of conduct on actual human beings, and inferring its probable effects on the human beings whom we may expect to exist in the near future.

3. Other thinkers of the evolutionist school [Sidgwick's phrase] suggest that the difficulties of utilitarian method might be avoided in a simpler way than Spencer's, by adopting as the *practically* ultimate goal and criterion of morality not happiness but the 'health' or 'efficiency' of the social organism. That is Leslie Stephen's view in *Science of Ethics*, chapter 9; it deserves careful examination. I understand Stephen to mean by 'health' the state of the social organism that tends to its preservation under the conditions of its existence, as they are known or capable of being predicted; and to mean the same by 'efficiency'. [Sidgwick explains the features of Stephen's writings that support this interpretation. Then:] The question, therefore, is this: If general *happiness* is admitted to be the really ultimate end in a system of morality, it is nevertheless reasonable to take *preservation of the social organism* as the practically ultimate 'scientific criterion' of moral rules?

I answer No, for two reasons. **(i)** I know no adequate grounds for supposing that if we aim exclusively at the preservation of the social organism we shall secure the maximum attainable happiness of its individual members. As far as I know, there's no limit to how different the happiness-levels could be of two social states that equally

tended to be preserved. As I pointed out in II/6.3 a large part of the pleasures that cultivated persons value most highly—*aesthetic pleasures*—are derived from acts and processes that don't significantly contribute to the preservation of •the individual's life or of •the social organism's life. Also, much refined morality is concerned with preventing pains that don't tend to the destruction of the individual or of society. I admit that the maintenance of preservative habits and sentiments is the most indispensable function of utilitarian morality and perhaps almost its only function in the earlier stages of moral development when living at all is a difficult task for human communities; but I don't infer from this that we should be content with merely securing survival for humanity generally, and should confine our efforts to promoting the increase of this security, instead of trying to make the secured existence more desirable.

(ii) I don't see why Stephen thinks that •the criterion of 'tendency to the preservation of the social organism' can be applied with greater precision than 'tendency to general happiness', even if the two ends coincide, and that •the former 'satisfies the conditions of a scientific criterion'. This probably *would* be the case if the sociology that we know were an actually constructed science and not merely the sketch of a possible future science; but Stephen himself has told us that sociology at present 'consists of nothing more than a collection of unverified guesses and vague generalisations, disguised under a more or less pretentious apparatus of quasi-scientific terminology'. I agree generally with this (though I wouldn't express it so strongly); and I don't see how a writer who holds this view can also maintain that the conception of 'social health', regarded as a criterion and standard of right conduct, is in any important degree more 'scientific' than the conception of 'general happiness'.

[In a further paragraph, Sidgwick says that his remarks about 'preservation of the social organism' apply also to 'development of the social organism'. A suggestion that the latter phrase might refer to 'definite coherent heterogeneity' generates an interesting footnote:] The increased heterogeneity that the development of modern industry has brought with it, in the form of a specialisation of industrial functions that tends to render the lives of individual workers narrow and monotonous, has usually been regarded by philanthropists as seriously unfelicitous, and as needing to be counteracted by a general diffusion of the intellectual culture now enjoyed by the few. If *that* came about it would tend to make the lives of different classes in the community less heterogeneous.

To sum up; I hold that in the present state of our knowledge the utilitarian can't possibly construct a morality from the ground up either for man •as he is (setting aside his morality) or for man •as he ought to be and will be. He must start with the existing social order, including the

existing moral code; and in deciding whether any divergence D from code C is to be recommended, he must mainly go by the immediate consequences of D on a society in which C is generally accepted. No doubt a thoughtful and well-instructed utilitarian may see dimly a certain way ahead, and what he sees may have some effect on his attitude towards existing morality.

- He may see certain evils threatening, which can't be warded off without adopting new and stricter views of duty in certain parts of life; and
- he may see a prospect of social changes that will make expedient or inevitable a relaxation of other parts of the moral code.

But if he keeps within the limits of scientific prevision, and doesn't stray into fanciful utopian conjectures, the form of society that he advocates won't differ much from the actual form of society, with its actually established code of moral rules and customary judgments about virtue and vice.

Chapter 5: The method of utilitarianism (continued)

1. Thus, a scientific utilitarian has a complex and balanced relation to the accepted morality of his age and country: common-sense morality is a machinery of rules, habits, and sentiments that is roughly and generally—but not precisely or completely—fitted for producing the greatest possible happiness for sentient beings generally; and the utilitarian has to *accept* it as the actually established machinery for achieving this goal, a machinery that can't be replaced all at once by a different one, but can only be gradually modified.

Now, how should the utilitarian *behave* in this situation?

Generally speaking, he will conform to the accepted morality and try to promote its development in others. Morality, considered as something accepted by human beings, isn't perfect; nothing in the human condition is perfect; from the human perspective *the universe* isn't perfect! But we should be much less concerned with •correcting and improving accepted morality than with •getting it to be obeyed. The utilitarian should entirely repudiate the attitude of rebellion

against the established morality as something purely external and conventional—the attitude that a reflective mind is always apt to acquire when it is first convinced that the established rules don't *stand to reason*. But of course he should also repudiate as superstitious the intuitional moralists' awe of established morality as an absolute or divine code.¹ Still, he will naturally contemplate it with reverence and wonder, as a marvelous product of nature, the result of long centuries of growth, showing in many parts the same fine adaptation of means to complex needs as the most elaborate structures of physical organisms exhibit; he will handle it with respectful delicacy as a mechanism, built out of the fluid element of opinions and dispositions, which provides indispensable help in the production of whatever human happiness is produced; a mechanism that no politicians or philosophers could create, yet without which the harder and coarser machinery of positive law [see Glossary] couldn't be permanently maintained, and the life of man would become—as Hobbes forcibly expresses it—'solitary, poor, nasty, brutish, and short'.

Still, this actual moral order *is* imperfect, so it's the utilitarian's duty to help in improving it (just as any law-abiding member of a modern civilised society sees law-reform as part of his political duty). How will he discover, at any given time and place, what changes in positive morality it would be practically expedient to try to introduce? Here we seem, after all, to be left with empirical hedonism as the only method that can ordinarily be used for the ultimate decision of such problems—at least until we have a real science of sociology. I'm not saying that the rudiments of sociological knowledge that we now have are of no practical value; because someone could suggest—and seriously well-meaning people some-

times *have* suggested—changes in morality that even our present scraps of knowledge lead us to regard as dangerous to the very existence of the social organism. But most such changes involve changes in positive law as well, because most of the fundamentally important moral rules are either directly or indirectly maintained by legal sanctions; and it would be going too far out of ethics and into politics to discuss such changes in the present book. When we are considering the utilitarian method of determining private duty, we'll have to deal mainly with rules that are supported by merely *moral* sanctions; and the question of whether to modify such a rule usually concerns the *well-being* of human society, not its very *existence*. So the utilitarian approach to this question comes down to comparing •the total amounts of pleasure and pain that can be expected from maintaining the rule in question with •the total amounts expectable from trying to introduce the proposed modification. This comparison must be of a rough and uncertain kind; we've already seen this, and it's important to bear it in mind, but we seem to have no substitute for it. I don't mean, of course, that each individual has to deal with such questions only through his own unaided judgment; there's a mass of traditional experience concerning the effects of conduct on happiness, and each individual can take this in either orally or from books; but the great formulae embodying this experience are mostly so indefinite, the proper range of their application is so uncertain, and the observations and inferences they are based on are so uncritical, that they continually need further empirical verification, especially as regards their applicability to any particular case.

So it's not surprising that some utilitarian thinkers •think that the task of hedonistic calculation that is thus set before

¹ I don't mean that this awe is incompatible with utilitarianism; I mean only that it mustn't be felt for any subordinate rules of conduct, but only for the supreme principle of acting with impartial concern for all elements of general happiness.

the utilitarian moralist is too big, and •propose to simplify it by marking off 'a large sphere of individual option and self-guidance', within which 'ethical dictation' doesn't apply. I admit that it's clearly expedient to draw a dividing line of this kind; but it seems to me that there's no simple general method of drawing it, and that the only way to draw it is through careful utilitarian calculation applied with varying results to people's real-life situations. To try to mark off the 'large sphere' by any such general formula as

'The individual is not responsible to society for the part of his conduct that concerns himself alone and others only with their free and undeceived consent'¹

seems to me to have no practical value. Why? Because the complex intertwining of interest and sympathy that connects people in a civilised community means that almost any significant loss of happiness by one person is likely to affect others—some quite considerably—without their consent. Mill says broadly that such secondary injury to others, if it is merely foreseen to be possible, is to be disregarded because of the advantages of allowing free development to individuality. I don't see how this can be justifiable from a utilitarian point of view. If the feared injury is great, and there's empirical evidence that it is very likely to ensue, I think that

the definite risk of evil from withdrawing the moral sanction

must outweigh

the indefinite possibility of loss through the repression of individuality in one particular direction.²

And there's another point: even if we *could* mark off the 'sphere of individual option and self-guidance' by some

simple and sweeping formula, a conscientious utilitarian will want *within* this sphere to guide take some account of how his actions affect the happiness of others; and the only methodical way to do this seems to be the empirical method that I discussed in Book II. Don't be too alarmed by this prospect: every sensible man [= 'every man whose feelings are in good order'] is commonly supposed to use pretty much this method in deciding on much of his conduct; it's assumed that within the limits that morality lays down he'll try to get as much happiness as he can for himself—and for others according to how they are related to him—by applying what he knows from his own experience and that of other men about the good and bad effects of actions. And that's how each man usually *does* think about

- what profession to choose for himself,
- what kind of education to choose for his children,
- whether to aim at marriage or remain single,
- whether to settle in town or country, in England or abroad,

and so on. I pointed out in III/14 that happiness isn't the only ultimate goal; knowledge, beauty, etc. are generally recognised as unquestionably desirable, and therefore often pursued with no thought of further consequences; but when the pursuit involves an apparent sacrifice of happiness in other ways, the practical question 'Should I continue the pursuit or abandon it?' is always decided by a rough application of the method of pure empirical hedonism. . . .

In determining the nature and importance of the various considerations that will come into play, the utilitarian art of morality [Sidgwick's phrase; for 'art' see Glossary] will get input from various sciences. It will learn from political economy

¹ This sentence, which is not an exact quotation, summarises the doctrine presented in the Introduction to Mill's *Liberty*.

² See Mill, *Liberty*, chapter 4. Mill's doctrine is certainly opposed to common sense: it would for example exclude from censure almost all forms of sexual immorality committed by unmarried and independent adults.

how the wealth of the community is likely to be affected by a general censure of usurers, or the routine approval of liberality in almsgiving; it will learn from the physiologist the probable effects on health of a general abstinence from alcohol, or of other restraint on appetite proposed in the name of temperance; it will learn from the experts in the relevant science *how far* knowledge is likely to be promoted by investigations—such as vivisection—that offend some prevalent moral or religious sentiment. But how far, in such-and-such circumstances, should the increase of wealth or of knowledge or good health be subordinated to other considerations? The only scientific method I know of for answering *that* is the method of empirical hedonism. Moralists used to label as ‘natural good’ everything that is intrinsically desirable apart from virtue or morality; and when men have been pursuing *that*, within the limits fixed by morality, the only method they have ever used is the one I have been describing. The utilitarian is merely performing—more consistently and systematically than ordinary men—the reasoning processes that are commonly accepted as appropriate to the questions that arise in the pursuit of natural good. What marks him off from the rest is that as a utilitarian he has to apply the same method to the criticism and correction of morality itself. The details of this criticism vary with the variations in human nature and circumstances: all I want here is to discuss the general points of view that a utilitarian critic must take. . . .

2. Let us first recall the distinction I presented in III/2 between **(i)** duty as commonly conceived—what a man is obliged to do—and **(ii)** praiseworthy or excellent conduct. In considering how utilitarianism relates to the moral judgments of common sense, I’ll start with **(i)** because it’s the more important and indispensable. [**(ii)** will be addressed on page 239.] That is, I’ll start with

the ensemble of rules imposed by common opinion in a society, forming a kind of unwritten legislation that supplements the law of the land and is and enforced by the penalties of social disfavour and contempt.

Because this legislation doesn’t come from a particular group of persons acting in a corporate capacity, it can’t be altered by any formal deliberations and decisions of the persons on whose consensus it rests. So any change in it must result from the private actions of individuals, whether or not they are determined by utilitarian considerations. The *practical* utilitarian problem is liable to be complicated by the conflicts and disagreements that occur between the moral opinions of different sections of almost any society; but at first I’ll attend only to rules of duty that are clearly supported by ‘common consent’. Let us suppose then that a utilitarian, after considering the consequences of rule R_1 , concludes that a it would be better for general happiness if R_1 were replaced by a different rule R_2 while the society remained unchanged in other respects. (It’s true of course that our forecast of social changes can’t easily be made clear enough to provide a basis for practice.) Let’s start with the case where R_2 differs from R_1 not only positively but **negatively**—it doesn’t merely go beyond and include R_1 but actually conflicts with it. [the ‘positive’ kind of moral amendment will be taken up on page 235.] Before the utilitarian can decide that it is right for him to support R_2 against R_1 by example and precept, he ought to estimate the force of certain disadvantages that are certain to accompany such innovations. They are of three kinds. . . .

(a) The happiness of the innovator and of his near and dear are a part of the end—universal happiness—at which he is aiming; so he must consider the importance to himself and them of the penalties of social disapproval that he will incur—not merely the immediate pain of this disap-

proval, but also its indirect effect making him less able to serve society and promote general happiness in other ways. The prospect of such pain and loss doesn't disqualify the innovation; . . . everything depends on the weight of those unpleasant effects, which can vary from •slight distrust and disfavour to •severe condemnation and social exclusion. It often seems that the •severest form of the moral penalty is imposed when an innovator attempts a moral reform *prematurely*, whereas if he had waited a few years he would have been let off with the •mildest. That is because a moral rule's hold on the general mind commonly begins to decay from the time that it is seen to be opposed to the calculations of expediency; and it may be better for the community as well as for the individual reformer if it isn't openly attacked until this process of decay has reached a certain point.

(b) More important are certain general reasons for doubting whether an apparent improvement really *will* have a beneficial effect on others. It's possible that the new rule R_2 , though it would be better than R_1 if it could get itself equally established, is less likely to be adopted, or if adopted less likely to be obeyed by the mass of the community in question. R_2 may be too subtle and refined, or too complex and elaborate; it may require a better intellect or more self-control than the average member of the community has, or an exceptional quality or balance of feelings. . . . Here as elsewhere in human affairs it is easier to pull down than to build up; easier to •weaken or destroy the restraining force that a habitually and generally obeyed moral rule has over men's minds than to •replace it a new restraining habit that isn't sustained by tradition and custom. So when the innovator by his own conduct sets an intrinsically •good example, the over-all effect may be •bad because its destructive operation proves to be more vigorous than its constructive. And the destructive effect can extend beyond

R_1 to all other rules. For just as the breaking of positive law has an inevitable tendency to encourage lawlessness generally, so the violation of any generally recognised moral rule seems to aid the forces that are always tending towards moral anarchy in any society.

(c) Any break with customary morality will have an effect on the innovator's own mind. The regulative habits and sentiments that each man has grown up with constitute an important force driving his will to act in ways that his reason would dictate. It's a natural auxiliary (so to speak) to reason in its conflict with seductive passions and appetites; and it may in practice be dangerous to weaken these auxiliaries. On the other hand, the habit of acting rationally seems to be the best of all habits, and a reasonable being should aim to bring all his impulses and sentiments into more and more perfect harmony with reason. Indeed, when a man has earnestly accepted a moral principle, those of his pre-existing regulative habits and sentiments that aren't in harmony with this principle tend naturally to decay and disappear; and it might be scarcely worthwhile to take them into account if it weren't for the support they get from the sympathy of others.

That support is a consideration of great importance. Each individual's moral impulses draw much of their effective force from the sympathy of other people. I don't mean merely this:

The pleasures and pains that each derives through sympathy from the moral likings and aversions of others are important not only as elements in the individual's happiness but also as motives to felicitous conduct.

I mean also this:

The direct sympathetic echo in each man of the judgments and sentiments of others concerning conduct sustains his own similar judgments and sentiments.

This twofold operation of sympathy makes it much easier

for most men to conform to a moral rule •established in their society than to one •made by themselves. And any act by which a man weakens the effect on himself of this general moral sympathy tends to that extent to make it harder for him to do his duty. •That is a *prima facie* reason against moral innovation, but here now is an extra reason in favour of it. As well as the intrinsic gain from the •particular change, there's the •general advantage of providing a striking example of consistent utilitarianism: a man gives a stronger proof of genuine conviction by opposing public opinion than he can by conforming to it. To get that effect, though, the non-conformity shouldn't favour the innovator's personal interests; for if it does, it will probably be attributed to egoistic motives, however plausible the utilitarian proof of its rightness may seem.

The considerations I have presented in this section will have different forces in different cases, and it's not worthwhile to attempt a general estimate of them. What we can say is that the general arguments that I have presented constitute an important *rational check* on negative or destructive utilitarian improvements on common-sense morality.

Let us turn now to innovations that are merely **positive** and supplementary, and consist in adding a new rule to those already established by common sense. [This positive/negative contrast was introduced on page 233.] The utilitarian's own observance of the new rule won't create any collision of methods. Every such rule is believed by him to be conducive to the common good, so he is merely giving a stricter interpretation to the general duty of universal benevolence, which common sense leaves loose and indeterminate. . . . And whatever it is right for *him* to do is obviously right for him to approve and recommend *others* to do in similar circumstances. But whether he should try to impose his new rule on others by condemning all who aren't prepared to adopt it—that's

a different question. Such conduct produces not only the immediate evil of the annoyance given to others but also the further danger of weakening—through the reaction provoked by this aggressive attitude—the general good effect of his moral example. What he decides about this will largely depend on what he thinks the chances are that his innovation will meet with support and sympathy from others.

Actually, much of the reform in popular morality that a consistent utilitarian tries to introduce will consist not so much in •establishing new rules (whether conflicting with the old or merely supplementing them) as in •enforcing old ones. There's always a considerable part of morality that receives formal respect and acceptance but isn't supported by any effective force of public opinion; and the different moralities of two societies may come less from disagreement about what rules the moral code should include than from differences in which of the rules they emphasise. The utilitarian's main task may be to get people to condemn more severely than they now do conduct that shows a lack of comprehensive sympathy or of public spirit. Such conduct often has the immediate effect of giving obvious pleasure to individuals while doing far greater harm more remotely and indirectly, and common sense is barely aware of the harm. So this conduct, even when it is agreed to be wrong, is very mildly treated by common opinion; especially when it is prompted by some impulse that isn't self-regarding. Such cases don't call for the promulgation of any new moral doctrine, but merely a bracing and sharpening of society's moral sentiments so as to bring them into harmony with •the greater breadth of view and •the more impartial concern for human happiness that characterise the utilitarian system.

3. . . . You may think that what utilitarianism and common sense are usually in conflict about is not •whether to introduce a new rule •or emphasise an old one• but rather

•whether exceptions should be allowed to rules that both sides accept as generally valid. While no-one doubts that it is, *generally speaking*, conducive to general happiness that men should

- be truthful,
- keep their promises,
- obey the law,
- satisfy the normal expectations of others,
- strictly control their malevolent impulses and sensual appetites,

some people think that an exclusive concern for pleasurable and painful consequences would often allow exceptions to rules that common sense imposes as absolute. Note, though, that admitting an exception on general grounds is merely establishing a more complex and intricate rule in place of one that is broader and simpler; for if it's conducive to the general good to admit an exception in one case, it will be equally so in all similar cases, and so introducing the exception is really instituting a new rule. Here is an example:

A utilitarian thinks it is right for him to answer falsely a question about how he has voted at a political election by secret ballot. He reasons that the utilitarian prohibition of falsehood is based on **(a)** the harm done by misleading particular individuals, and **(b)** the tendency of false statements to lessen men's confidence in one another's *assertions*; and that in this exceptional case it is **(a)** expedient that the questioner should be misled; while **(b)** in tending to produce a general distrust of all *assertions about how a man has voted*, the falsehood only furthers the end for which voting has been made secret.

If these reasons are valid for one person they are valid for everyone. In fact, they show the expediency of a new general rule concerning truth and falsehood, more complicated than

the old one—a rule that a utilitarian should desire to be universally obeyed.

Some kinds of moral innovation are unlikely to occur often—e.g. utilitarian reasoning leads a man to take part in a political revolution, or to support a public measure that conflicts with what common sense regards as justice or good faith. But in such cases—rare or not—a rational utilitarian will usually proceed on general principles that he would like to be applied by anyone in similar circumstances.

Utilitarianism seems to allow another fundamentally different kind of exception—one in which the agent doesn't think it expedient that the rule on which he himself acts should be universally adopted, and yet maintains that his individual act is right by utilitarian standards. . . . Just as a prudent physician in giving rules of diet recommends an occasional deviation from them, as better for the health of the body than absolute regularity, the same may hold for some rules of social behaviour. It might be that the general observance of a certain rule is necessary for the community's well-being though a certain amount of non-observance is advantageous rather than harmful.

Here we seem to be in conflict with Kant's fundamental principle that a right action must be one of which the agent could 'will the maxim to be a universal law' (see above pages 98 and 188). But, as I pointed out in III/7.3, in the particular case of veracity we must regard the maxim that the Kantian principle is supposed to test to *include* the qualification ' . . . if the agent believes that this action won't be widely imitated'. Kant's principle, in the only version of it that I have accepted as self-evident, means only that

If an act is right for some individual, it must be right on general grounds, and therefore right for some *class* of persons;

so it can't prevent us from defining this class as 'those who

believe that the act in question will remain an exceptional one'. If this belief turns out to be erroneous, serious harm may result; but that's true also of many other utilitarian deductions. And it's easy to find examples of conduct that common sense permits solely because we're sure it won't be widely imitated—celibacy, for example. A universal refusal to propagate the human species would be the greatest conceivable crime from a utilitarian point of view—i.e. according to the commonly accepted belief in the superiority of human happiness to that of other animals—so that Kant's principle, if not qualified in the way I have described, would make it a crime in anyone to choose celibacy as the state most conducive to his own happiness. But common sense (in the present age at least) regards such a preference as within the limits of right conduct, because there's no fear that population won't be sufficiently kept up (in fact the tendency to propagate is thought to exist in excess!). [The 'belief in the superiority of human happiness to that of other animals' is flatly irrelevant to the morality of closing out the human race; at this point Sidgwick seem to have blundered. For a discussion of moral issues that *do* arise regarding the continuation of *Homo sapiens* you might visit Bennett, 'On Maximizing Happiness' at <http://www.earlymoderntexts.com/jfb/maxhap.pdf>]

In this case we are relying on the average strength of a •non-moral impulse; but there seems to be no formal or universal reason why the same procedure shouldn't be applied by utilitarians in reliance on an existing •moral sentiment. The result would be an odd discrepancy between utilitarianism and common sense morality: the very firmness with which the latter is established would be the utilitarian ground for relieving the individual of his obligations. We're supposed to see that general happiness will be enhanced. . . .by a slight admixture of irregularity along with a general observance of received rules; and hence to justify

the irregular conduct of a few individuals, on the ground that the supply of regular conduct from other members of the community may reasonably be expected to be adequate.

[Sidgwick goes into this at wearying length. He concludes that a conscientious person will almost never be sure *enough* that his rule-breach won't weaken the rule's hold on people in general to think he is morally entitled to break the rule just this once. After a page of this, he continues:]

So it seems to me that the cases in which practical doubts arise about whether utilitarian principles allow exceptions to ordinary rules will mostly be the ones I discussed early in this section [page 235], where the exceptions are claimed

- not for a few individuals merely because they are few, but either
- for persons generally under exceptional circumstances, or
- for a class of persons defined by exceptional qualities of intellect, temperament, or character.

[Don't be misled by the difference between 'persons' and 'class of persons'. What matters is the difference between •being in an exceptional situation and •having exceptional qualities.] In such cases the utilitarian may be sure that in a community of enlightened utilitarians these grounds for exceptional ethical treatment would be regarded as valid; but he may doubt whether the more refined and complicated rule that recognises such exceptions is adapted for the community in which he is actually living; and may suspect that the attempt to introduce the new rule will do more •harm by weakening current morality than •good by improving its quality. . . . He should consider carefully how likely his advice or example are to influence persons to whom they would be dangerous; and it's clear that the answer to this will depend largely on how *publicly* he is going to offer his advice or example. On utilitarian principles it can be right to do and privately recommend. . . .something that it

would not be right to advocate openly; it may be right to teach openly to one set of persons what it would be wrong to teach to others; it may conceivably be right to do in comparative secrecy something that it would be wrong to do in the face of the world. . . . These conclusions are all paradoxical;¹ there's no doubt that the moral consciousness of a plain man rejects the notion of an esoteric morality differing from the morality that is taught to the people; and it would be commonly agreed that an action that would be bad if done openly isn't made good by secrecy. There are indeed strong utilitarian reasons for generally maintaining this latter common opinion. Here are two of them. (i) It is obviously advantageous that acts that it's expedient to repress by social disapproval should become known, as otherwise the disapproval can't operate; so that it seems inexpedient to give any moral encouragement to men's natural disposition to conceal their wrong doings. (ii) Such concealment would usually do significant harm to the agent's habits of veracity. So the utilitarian conclusion, carefully stated, seems to be this:

The opinion that *secrecy can make right an action that wouldn't otherwise be so* should itself be kept comparatively secret;

and similarly it seems expedient that the doctrine that *esoteric morality is expedient* should itself be kept esoteric. If this concealment is hard to maintain, it may be desirable that common sense should repudiate completely the doctrines that it's expedient to confine to an enlightened few. And thus a utilitarian may reasonably desire *on utilitarian principles* that some of his conclusions should be rejected by mankind generally; or even that ordinary folk should keep their distance from his system as a whole because the inevitable indefiniteness and complexity of its calculations

make it likely to lead to bad results in their hands. In an ideal community of enlightened utilitarians (I repeat) this swarm of perplexities and paradoxes would vanish, because in such a society no-one can have any reason to think that anyone else will act on moral principles different from his. And of course any enlightened utilitarian must want this state of affairs to come about: all conflict of moral opinion is to some extent bad because it tends to lessen the power with which morality resists seductive impulses. Still such conflict may be a necessary evil in civilised communities as they actually are, with so many different levels of intellectual and moral development.

So I have been led to discuss the question that I set aside near the start of section 2:

How should utilitarianism handle the fact of divergent moral opinions among different members of the same society?

It has become plain that although two different kinds of conduct can't both be right under the same circumstances, two conflicting opinions about the rightness of conduct may both be expedient. It may be best for general happiness that A should perform a certain action while B, C and D blame it. The utilitarian can't really join in the disapproval, but he may think it best to leave it unshaken, while also thinking it would be right for him, if placed in the supposed circumstances, to perform the action that is generally disapproved. And so it may be best over-all that there should be conflicting codes of morality in a given society at a certain stage of its development. And the reason for holding that

(a) common-sense morality roughly coincides with the utilitarian code that is appropriate for men as now constituted

¹ In particular cases, however, common sense seems to admit them to a certain extent. It would commonly be thought wrong to express in public speeches disturbing religious or political opinions that it's all right to publish in books.

is also a reason for holding that

- (b) these divergent moral codes are also appropriate for men as now constituted, and are needed to supplement and qualify the morality of common sense.

The reason for (a) involves the probable origin of the moral sense and its flexible adjustment to the varying conditions of human life; and the divergent moral codes in (b) are also part of man's complex adjustment to his circumstances.

Paradoxical as it seems to be, this doctrine sometimes seems to be •implicitly accepted by common sense or at least to be •required to make common sense self-consistent. Concerning rebellions, for example. It is commonly thought •that these abrupt breaches of order are **sometimes** morally necessary, and also •that they ought **always** to be vigorously resisted, and in case of failure punished by extreme penalties, at least for the ring-leaders, because otherwise rebellions would be attempted in circumstances where there was no sufficient justification for them. But it seems evident that—given the actual condition of men's moral sentiments—this vigorous repression needs to be backed by a strong body of opinion condemning the rebels as •wrong and not merely •mistaken in their expectation of success. For similar reasons it might be expedient on the whole that certain special relaxations of certain moral rules should continue to exist in certain professions and sections of society, while continuing to be disapproved of by the rest of the society. But the evils that are bound to arise from this permanent conflict of opinion are so grave that an enlightened utilitarian will probably try to remove it in most cases, either

- by openly maintaining the need for the ordinary moral rule to be relaxed in those special circumstances, or
- by trying to get the ordinary rule recognised and enforced by all conscientious persons in the section of society where breaking it has become habitual.

It's likely that in most cases he will take the latter approach, because such rules are usually found on examination to have been relaxed for the convenience of individuals rather than the good of the community at large.

4. Finally, let us consider how utilitarianism relates to the part of common morality that extends beyond the range of strict duty—i.e. to the ideal of character and conduct that a given community at a given time admires and praises as the sum of excellences or perfections. [This is item (ii) of the pair announced at the start of section 2 on page 233.] This distinction between strict duty and excellence seems not to be properly admissible in utilitarianism (except for excellences that aren't wholly and directly under the control of the will; we should distinguish •conduct displaying these from •the doing of duty, which can always be done at any moment); because a utilitarian must hold that it's always wrong for a man to do anything except what he believes to be most conducive to universal happiness. Still, it seems to be practically expedient—and therefore indirectly reasonable on utilitarian principles—to distinguish conduct that is *praiseworthy and admirable* from conduct that is merely *right*, even when all the conduct in question is strictly voluntary. Why? Well, it's natural for us to compare an individual's character or conduct not with •our highest ideal—utilitarian or otherwise—but with a certain •average standard, and to admire anything that rises above that standard; and it seems to be conducive to general happiness that such natural sentiments of admiration should be encouraged and developed. To come up with the best performance of duty that is currently possible for it, human nature seems to require the double stimulus of blame *and* praise from others; so that the 'social sanction' would be less effective if it became purely penal—i.e. included the blame and left out the praise. And utilitarianism itself is opposed to relying solely on blame, because remorse and disapproval

are painful. . . . But there is still a reasonable place for the aesthetic phase of morality: we may properly admire and praise where it would be inexpedient to judge and condemn. So it is reasonable for a utilitarian to praise conduct that is better—contributes more to happiness—than what an average man would do under the same circumstances; not forgetting that the lower limit to praiseworthiness should be relative to the state of moral progress of people in the country concerned, and that it is desirable to make continual efforts to raise this standard. . . .

How does the utilitarian ideal of character compare with the virtues and other excellences recognised by common sense? Well, there's a general coincidence between the two that Hume and others have emphasised. Any quality that has ever been praised as excellent by mankind generally can be shown to have some marked felicific effect, and to be. . . .obviously conducive to general happiness. But it doesn't follow that society always fosters and encourages such qualities in the proportion that a utilitarian would desire; in fact, we often see societies where some useful qualities are unduly neglected while others are over-prized and even admired though they exist in such excess as to become over-all infelicitous. The consistent utilitarian may therefore find it necessary to correct the prevalent moral ideal; and he won't run into utilitarian restrictions on correction of **ideals** as we found in correcting commonly accepted rules of **duty**. For the common-sense notions of excellences of conduct that go beyond the range of strict duty are generally too vague to offer any definite resistance to a utilitarian interpretation of their scope; a man can teach and act on the basis of such an interpretation without risking a harmful conflict with common sense—especially given that the ideal of moral excellence varies much more widely than the code of strict duty does in the same community. A man who

- at a time and place where excessive asceticism is praised sets an example of enjoying harmless bodily pleasures, or
- in social circles where useless daring is admired prefers to exhibit and commend caution and discretion,

at the worst misses some praise that he might otherwise have earned, and is thought a little dull or unambitious; he doesn't come into any obvious conflict with common opinion. An enlightened utilitarian is likely to lay less stress on the cultivation of •negative virtues—tendencies to restrict and refrain—that loom large in the common sense ideal of character; and to set more value on qualities of mind that are the direct source of •positive pleasure to the agent or to others, some of which common sense scarcely recognises as excellences. But he won't carry this innovation so far as to get himself generally condemned. For an enlightened utilitarian can't ignore the fundamental importance of the restrictive and repressive virtues, and can't think they are *now* so well developed in ordinary men that there's no need to encourage them by moral admiration. . . . Under most circumstances, indeed, a man who earnestly and successfully tries to bring about the utilitarian ideal, however he may deviate from the commonly accepted notion of a perfect character, is likely to win enough recognition and praise from common sense. Here is why:

Whether or not the whole of morality has sprung from the root of sympathy, it's certain that self-love and sympathy combined are strong enough in average men to dispose them to grateful admiration of any exceptional efforts to promote the common good, even if these efforts take a novel form. Common sense nearly always reacts well to any exhibition of more extended sympathy or more fervent public spirit than

is ordinarily shown, and to any attempt to develop these qualities in others—provided of course that these impulses are accompanied with adequate knowledge of actual circumstances and insight into the relation of means to ends, and that they don't conflict with any recognised rules of duty. And it's principally in this direction that the recent spread of utilitarianism has positively modified the ideal of our society, and is likely to modify it further. That is why utilitarians are apt to stress social and political activity of all kinds, and why utilitarian ethics have always tended to pass over into politics. Someone

[That concludes Book IV. The remaining chapter, not numbered by the author, looks across the entire work and doesn't belong to Book IV in particular.]

who values conduct in proportion to its felicitous consequences will naturally value •effective beneficence in public affairs more than even the purest •manifestation of virtue in private life; whereas on the other hand an intuitionist (though no doubt vaguely recognising that a man ought to do all the good he can in public affairs) still commonly holds that virtue may be as fully and admirably exhibited on a small as on a large scale. A sincere utilitarian, therefore, is likely to be an eager politician. What principles should guide his political activity? Searching for an answer to that is a task outside the scope of this treatise.

Concluding chapter: The mutual relations of the three methods

1. Throughout most of this work I have been employed in examining three methods of determining right conduct—methods that are more or less vaguely combined in the practical reasonings of ordinary men, though I have tried to expound them as separately as possible. I shan't attempt here a complete synthesis of these methods, but I shouldn't conclude the analysis of them without some discussion of their relations to one another. I have indeed already found it expedient to do a good deal of this while examining the separate methods. I have directly or indirectly examined quite fully the relations between the intuitional and utilitarian methods. I have shown that the common antithesis between intuitionists and utilitarians must be

entirely discarded, because abstract moral principles that we can admit to be really self-evident are not only *compatible with* a utilitarian system but even seem to be needed as a rational *basis for* such a system. Example: the essence of justice or equity (insofar as it is clear and certain) is that different individuals are not to be treated differently except on grounds that apply universally; and such grounds are supplied by the principle of universal benevolence that tells each man that the happiness of all others is as worthy a goal as his own; while other time-honoured virtues seem to be •special manifestations of impartial benevolence in various special circumstances, or •habits and dispositions that are needed for the maintenance of prudent or beneficent

behaviour under the seductive force of various non-rational impulses. There are rules that our common moral sense seems at first to pronounce as absolutely binding, but it has turned out that these results are really subordinate to the fundamental principles on which utilitarianism is based—this being something we learn about them by careful and systematic reflection on common sense itself, as expressed in the habitual moral judgments of ordinary men.

This way of looking at particular virtues and duties is strongly supported by from a comparative study of the history of morality. The variations in the moral codes of different societies at different stages largely correspond to differences in the actual or believed tendencies of certain kinds of conduct to promote the general happiness of different portions of the human race; and the most probable conjectures about the pre-historic condition origin of the moral faculty seem to be entirely in harmony with this view. The **(i)** results of special utilitarian calculations won't of course always agree in all the details with **(ii)** our particular moral sentiments and unreasoned judgments, and we may often find it hard *in practice* to balance **(i)** against the more general utilitarian reasons for obeying **(ii)**; but there seems to be no longer any *theoretical* perplexity about the principles for determining social duty.

Regarding the two species of hedonism that I have distinguished as 'universalistic' and 'egoistic'—how are they related to one another? In chapter 2 I discussed the rational process (called by a stretch of language a 'proof') by which someone who holds it reasonable to aim at his own greatest happiness may be brought to take universal happiness instead as his ultimate standard of right conduct. And we have seen that this process doesn't work unless the egoist affirms, implicitly or explicitly, that his own greatest happiness is not merely •the rational ultimate end for himself but •a part of

universal good; and he can avoid the 'proof' of utilitarianism by declining to affirm this. Common sense won't let him deny that the distinction between himself and any other person is real and fundamental; so it puts him in a position to think:

'I am concerned with the quality of my existence as an individual in a fundamentally important sense in which I am not concerned with the quality of the existence of anyone else';

and I don't see how it can be proved that this distinction ought not to be taken as fundamental in fixing the ultimate goal of an individual's rational action. Notice that most utilitarians, however much they have wanted to convince everyone of the reasonableness of aiming at happiness generally, haven't commonly tried to do this through any logical transition from the •egoistic to the •universalistic principle. They have relied almost entirely on the pleasures gained or pains avoided by anyone who conforms to the utilitarian rules. Indeed, if an egoist isn't moved by what I have called *proof*, the only way of arguing him into aiming at everyone's happiness to show that this gives him his best chance of greatest happiness for himself. And even if he admits that the principle of rational benevolence is self-evident, he may still hold •that it is irrational for him to sacrifice his own happiness to any other end; and •that therefore

If morality is to be made completely rational the harmony between the maxim of prudence and the maxim of rational benevolence must be somehow demonstrated.

I have said before that this latter view seems to be what common sense holds; and I hold it too. So we should examine how far and in what way the required demonstration can be effected.

2. Some of that investigation was done in II/5, where it appeared that while

in any tolerable state of society the virtuous agent's best chance of achieving •his own greatest possible happiness in the long run is likely to come from his •exercising the social virtues,

there's no empirical evidence that the two will *always* coincide and *completely* coincide; and that indeed the more carefully we examine how the different sanctions—legal, social, and conscientious—operate in the actual conditions of human life, the harder it is to believe that they can always produce this coincidence of happiness with social virtue. This will merely motivate a convinced utilitarian to try to alter the actual conditions of human life; and it would be a valuable contribution to the actual happiness of mankind if we could in any society

- fine-tune the machine of law,
- stimulate and direct the common awards of praise and blame, and
- develop and train the moral sense of the members of the community

in such a way as to make it clearly prudent for every individual to do everything he can for the general good. But our present topic is not •what a consistent utilitarian will try to bring about in the future but •what a consistent egoist should do in the present! And it must be admitted that in the present

state of the world •egoism has a better chance of coinciding with •common-sense morality than with •utilitarian morality; because—as we have seen—utilitarianism is more rigid than common sense in demanding that the agent sacrifice his private interests when they are incompatible with the greatest happiness of the greatest number. . . .

3. But some utilitarian writers¹ seem to think that we can be led to see that the good of each coincides with the good of all by thinking hard and well about the paramount importance of sympathy as an element of human happiness. In opposing this view, I am as far as possible from any wish to depreciate the value of sympathy as a source of happiness even to human beings as at present constituted. [Sidgwick develops this thought in an enormous aside, or subordinate clause. He resumes what he was starting to say here, namely that the 'utilitarian writers' in question are wrong, in the paragraph starting '**But allowing all this.** . . .' on page 244. Notice that after the first *two words* of the resumption, he has a footnote in which he ducks back into the aside!] Indeed I hold that the pleasures and pains of sympathy constitute a great part of the internal reward for social virtue and punishment for social misconduct that I roughly described in II/5 as due to the moral sentiments. When I look into my own consciousness, I can to some extent distinguish sympathetic feelings from strictly moral ones, but I can't say precisely in what proportion the two are combined. For instance, it seems that I can distinguish •the 'sense of the ignobility

¹ See Mill, *Utilitarianism*, chapter 3; though the argument there is hard to follow because it mixes up three different objects of inquiry: **(1)** the actual effect of sympathy in inducing conformity to the rules of utilitarian ethics, **(2)** the effect in this direction that it's likely to have in the future, **(3)** the value of sympathetic pleasures and pains as estimated by an enlightened egoist. Mill didn't clearly separate **(1)** from **(3)**, because of his psychological doctrine that each person's own pleasure is the sole object of his desires. But if my refutation of this doctrine in I/4.3 is valid, we have to distinguish two ways in which sympathy operates: it *generates sympathetic pleasures and pains*, which have to be taken into account in the calculations of egoistic hedonism; but it may also *cause impulses to altruistic action* the force of which is quite out of proportion to the sympathetic pleasure (or relief from pain) that such actions seem likely to bring to the agent. So that even if the average man did ever reach such a pitch of sympathetic development that he never felt prompted to sacrifice the general good to his own, this still doesn't prove that it is egoistically reasonable for him to behave in this way.

of egoism'. . . from •the jolt of sympathetic discomfort that accompanies the conscious choice of my own pleasure at the expense of pain or loss to others; but I can't determine what force the former sentiment would have if actually separated from the latter, and I'm inclined to think that the two kinds of feeling are very differently combined in different individuals. It may be that in the development of the moral consciousness of mankind and of individual men, a general law operates concerning the relative proportions of these two elements; for it seems that at a certain stage in this development the mind is more susceptible to emotions connected with •abstract moral ideas and rules presented as absolute; whereas before entering this stage and after emerging from it •the feelings that belong to personal relations are stronger.¹ Certainly in a utilitarian's mind sympathy tends to loom large in all instinctive moral feelings that refer to social conduct; just as in his view the rational basis for the moral impulse must ultimately lie in some pleasure won or pain saved for himself or for others; so that he never has to sacrifice himself to an impersonal law but always in the interests of some beings with whom he has some degree of fellow-feeling.

And I would go further and maintain—simply on empirical grounds—that enlightened self-interest would direct most men to give sympathetic feelings a larger role in their lives than it commonly does now. There's no denying the effectiveness of Butler's famous argument against the vulgar antithesis between •self-love and •benevolence; and it isn't much of an exaggeration to say that amid all the profuse

waste of the means of happiness that men commit there's no imprudence more flagrant than that of *selfishness* in the ordinary sense of that word—the excessive concentration on one's own happiness that makes it impossible for one to feel any strong interest in the pleasures and pains of others. The perpetual prominence of *self* that comes from this tends to deprive all enjoyments of their keenness and zest, and quickly produces satiety and boredom; the selfish man misses the sense of elevation and enlargement given by wide interests; he misses the more secure and serene satisfaction that continually accompanies activities that are directed towards goals that are more stable than an individual's happiness can be; he misses the special rich sweetness, coming from a complex reverberation of sympathy, that is always found in services rendered to those whom we love and who are grateful. He is made to feel in a thousand different ways, according to the level that his nature has reached, the discord between the rhythms of his own life and of those of the larger life of which his own is only an insignificant fraction.

But allowing² all this, it still seems to me to be certain—so far as any conclusion based on hedonistic comparison *can* be certain—that the utmost development in the strength and scope of sympathies that is now possible to any but a very few exceptional persons would not cause utilitarian duty coincide perfectly with self-interest. Here it seems to me that what I said in II/5.4 to show the insufficiency of punishment by conscience applies equally, *mutatis mutandis*,

¹ I do not mean to imply that the process of change is merely circular. In the earlier period sympathy is narrower, simpler, and more presentative; in the later it is more extensive, complex, and representative. [He means that first the sympathy is just a self-contained feeling, whereas later it comes to mean or be *about* something; first it merely *presents* itself, and later it *represents* something else.]

² I don't think that we *should* allow what I have been saying as universally true. A few thoroughly selfish persons at least *seem* to be happier than most of the unselfish; and there are other exceptional natures whose chief happiness seems to come from activities which, though disinterested, are directed towards ends other than human happiness.

to punishment by sympathy. Suppose a man finds that a concern for the general good—utilitarian duty—demands that he sacrifice his life or incur an extreme risk of doing so. There may be one or two people who are so dear to him that the rest of a life saved by sacrificing their happiness to his own would be worthless to him from an egoistic point of view. But it is doubtful whether many men, ‘sitting down in a cool hour’ to make the estimate, would affirm even this. [He presumably means: ‘could truthfully say: “There are people I love so much that if I sacrificed their interests to my own, the rest of my life wouldn’t be worth living”.’] And of course the particular portion of the general happiness for which one is called on to sacrifice one’s own may easily be the happiness of folk one isn’t especially fond of. It is normal for us to limit our keenest and strongest sympathy to a very small circle of people; and a result of that is that the *development* of sympathy may increase the weight thrown into the scale *against* utilitarian duty. Very few people, however strongly and widely sympathetic, feel for the pleasures and pains of mankind generally a degree of sympathy comparable with their concern for wife or children, or lover, or intimate friend; and if any training of the affections is at present possible that would materially alter this proportion in the general distribution of our sympathy, it doesn’t look as though such training would be on the whole felicitous (see chapter 3.3). Thus, when utilitarian duty calls on us to sacrifice to the general good not only our own pleasures but the happiness of those we love, the very sanction on which utilitarianism most relies—namely sympathy—must act powerfully in opposition to its precepts.

The cases I have been discussing are exceptional, but they do decide the abstract question. Even setting them aside, it seems that the conduct by which a man would most fully reap the rewards of sympathy. . . . will often be different

from the conduct dictated by a sincere desire to promote general happiness. The relief of distress is an important part of utilitarian duty; but the state of the beneficiary is painful, so that sympathy with him seems to be a source of pain rather than pleasure; how much pain depends on how intense the sympathy is. It’s probably true in general that in the relief of distress other elements of the complex pleasure of benevolence decidedly outweigh this sympathetic pain; because

- the welling-up of pity is itself pleasurable, and
- we commonly feel the improvement of the sufferer’s state that we have produced more keenly than we do his pain that was caused in some other way, and
- there’s further the pleasure that we get from his gratitude, and
- there’s pleasure that is a normal upshot of activity directed under a strong impulse towards a permanently valued end.

Still, when the sufferer’s distress is bitter and continued, and we can only partly relieve it by all our efforts, the benefactor’s sympathetic discomfort must be considerable; and the work of combating misery, though it does have some elevated happiness, will be much less happy over-all than many other forms of activity; yet it may be just *this* work that duty seems to summon us to. Or a man might find that he can best promote general happiness by working

- in solitude for ends that he never hopes to see achieved, or
- chiefly for people for whom he can’t feel much affection, or
- on projects that must alienate or grieve those he loves best, or
- on projects that require him to dispense with the most intimate of human ties.

There seems to be no end to the ways in which the dictates of rational benevolence—which as a utilitarian he is obliged absolutely to obey—can conflict with that indulgence of kind affections that Shaftesbury and his followers so persuasively exhibit as its own reward.

4. So it seems that we must conclude, from the arguments in III/5 supplemented by the discussion just completed, that the inseparable connection between •utilitarian duty and •the greatest happiness of the individual who conforms to it cannot be satisfactorily demonstrated on empirical grounds. This has led other utilitarian writers to prefer to throw the weight of duty on the *religious sanction*; and this procedure has been partly adopted—even by some of those who have chiefly dwelt on sympathy as a motive. From this point of view the utilitarian code is conceived as the law of God, who is to be regarded as having commanded men to promote general happiness, and announced that he will reward those who obey his commands and punish those who don't. [In this next sentence, the word 'feel' is Sidgwick's.] It's clear that if we feel convinced that an omnipotent being has somehow signified such commands and announcements, a rational egoist can't need any further inducement to shape his life on utilitarian principles. The only question is *How did he get this conviction?* This is commonly thought to be either by •supernatural revelation or by •the natural exercise of reason, or •in both ways. As regards revelation, nearly all the moralists who hold that God has disclosed his law either •to special individuals in past ages who left a written record of what was revealed to them, or •to an unbroken succession of persons appointed in a particular manner, or •to religious persons generally in some supernatural way, think that what is thus revealed is not the utilitarian code but rather the rules of common-sense morality with some special modifications and additions. But Mill was right to stress

that utilitarianism, being more rigorous than common sense in demanding the sacrifice of the individual's happiness to that of mankind generally, it is strictly in line with the most characteristic teaching of Christianity. There's no need for me to discuss the precise relation of different revelational codes to utilitarianism; it would be going beyond the limits of my topic to go into why a divine origin has been attributed to them.

Given the belief that a knowledge of God's law can be attained by the reason, ethics and theology seem to be so closely connected that we can't draw a sharp line between them. As we saw in III/1.2 and chapter 2.1, it has been widely maintained that the relation of moral rules to a divine lawgiver is implicitly recognised in the act of thought by which we discern these rules to be binding. And no doubt the terms (such as 'moral obligation') that we commonly use in speaking of these rules naturally suggest •legal sanctions and thus •a sovereign by whom these are announced and enforced. Indeed many thinkers since Locke have said that the only meaning for the terms 'right', 'duty', etc. is that of a rule imposed by a lawgiver. But this view seems contrary to common sense; perhaps the easiest way to show this (see I/3.2) is to point out that the divine lawgiver is himself thought of as a moral agent, i.e. as prescribing what is right and designing what is good. It's clear that in *this* thought, at least, the notions 'right' and 'good' are used without any reference to a superior lawgiver; and religious persons seem to hold that the words are used here in a sense not essentially different from their ordinary meanings. Still, although common sense does not regard moral rules as being *merely* the commands of an omnipotent being who will reward and punish men according as they obey or violate them, it certainly holds that this is a true though partial view of them, and perhaps that it can be known intuitively.

If then reflection leads us to conclude that common sense's moral principles are to be systematised as subordinate to the pre-eminently certain and unshakable intuition which stands as the first principle of utilitarianism, then of course it will be the utilitarian code that we'll believe the divine sanctions to be attached to.

Or we might argue thus. If we are to conceive of God as acting for some end—as all theologians agree that he does—we must conceive that end to be •universal good, and if utilitarians are right •universal happiness; and we can't suppose that in a morally governed world it can be prudent for anyone to act in conscious opposition to what we believe to be the divine design. Hence if after calculating the consequences of two alternatives of conduct we choose the one that seems likely to be less conducive to happiness generally, we'll be acting in a manner that we must expect to suffer for.

It has been objected against this that we can see that the happiness of sentient beings is so imperfectly attained in the actual world, and is mixed with so much pain and misery, that we can't really think that universal happiness is God's end unless we admit that he isn't omnipotent. No doubt the assertion that God is omnipotent will need to be understood with some limitation, but perhaps with no greater limitation than has always been implicitly admitted by thoughtful theologians, who seem always to have accepted that there are things that God can't do, e.g. change the past. And if our knowledge of the universe were complete, perhaps we would see that the quantum of happiness ultimately

attained in it is as great as could be achieved without doing something that we would then see to be just as inconceivable and absurd as changing the past. But this is a line of thought for the theologian to develop. What I want to stress is that apparently none of the other ordinary interpretations of 'good' does any better than utilitarianism in how good it implies the actual universe to be. The wonderful perfections of work that we admire in the physical world are all mingled with imperfection and liable to destruction and decay; and similarly in the world of human conduct virtue is at least as much balanced by vice as happiness is by misery. So that, if the ethical reasoning that led us to interpret ultimate good as happiness is sound, there seems no argument from natural theology [see Glossary] to set against it.

5. So if we can assume the existence of a being such as God is said (by the theologians) to be, it seems that utilitarians are entitled to infer that there are divine rewards (or punishments) for obeying (or violating) the code of social duty that arises out of utilitarianism; and of course these would make it always in everyone's interests to promote universal happiness to the best of his knowledge. But what *ethical* grounds are there for the assumption of God's existence? The answer to this will settle the question of whether ethical science can stand on its own feet or whether it is forced to borrow a fundamental and indispensable premise from theology or some similar source.¹ In order to approach this question fairly, let us reflect on the clearest and most certain of our moral intuitions. I find that I undoubtedly seem to perceive—as clearly and certainly as I see any axiom in

¹ If we are simply considering ethics as a possible independent science, the fundamental premise whose validity we are now examining doesn't have to have a theistic form. And it apparently hasn't always taken that form in the support that positive religion has given to morality. In the Buddhist creed this notion of the rewards for right conduct seems to have been developed in a far more elaborate and systematic manner than it has in any branch of Christianity. But enlightened Buddhists see these rewards as distributed not by •the will of a supreme person but by •the natural operation of an impersonal law.

arithmetic or geometry—that it is ‘right’ and ‘reasonable’ for me to treat others as I think that I ought to be treated under similar conditions, and to do what I believe to be ultimately conducive to universal good or happiness. But I can’t find inseparably connected with this conviction, and similarly attainable by mere reflective intuition, any cognition that there actually is a supreme being who will adequately¹ reward me for obeying these rules of duty, or punish me for violating them.² Or—omitting the strictly theological element of the proposition—I can report that I do not find in my moral consciousness any intuition, claiming to be clear and certain, that the performance of duty will be adequately rewarded and its violation punished. I do indeed feel a *desire*—apparently inseparable from the moral sentiments—for this to be the case not only for me but for everyone; but the mere existence of the desire doesn’t go far to establish the probability of its fulfilment! I also judge that in a certain sense it *ought* to be the case that rewards and punishments are distributed according to people’s deserts; but in this judgment ‘ought’ is not used in a strictly ethical meaning; it only expresses our practical reason’s feeling that it can’t be made consistent with itself unless it proves or postulates this connection between virtue and self-interest. Denying this would force us to admit an ultimate and basic 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, shown in these contradictory judgments, is after all illusory.

I do not mean that if we gave up the hope of resolving this basic contradiction through a legitimately obtained conclusion or postulate about the world’s moral order, it would become reasonable for us to abandon morality altogether; but it seems that we would have to abandon the idea of rationalising it completely. [He means: ‘give up the idea of capturing morality in a consistent and comprehensive set of general propositions or rules’.] We would no doubt still feel a desire for the general observance of rules conducive to general happiness, being led to this not only by self-interest but also by sympathy and sentiments protective of social well-being that we had learned through education; and practical reason would still impel us decisively to the performance of duty in ordinary cases where what is recognised as duty is in harmony with self-interest properly understood. But in the rarer cases where we find a 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 by which of two groups of *non-rational* impulses had more force.

So we have this:

•**The harmony of duty and self-interest** is a hypothesis that is required if we are to avoid a basic contradiction in one chief part of our thought.

So the question arises:

•Is the above fact a sufficient reason for accepting this hypothesis?

This is a profoundly difficult and controverted question. The discussion of it belongs to a treatise on general philosophy rather than to a work on the methods of ethics, because it

¹ Remember that by ‘adequate’ I mean ‘sufficient to make it the agent’s interest to promote universal good’, not necessarily ‘proportional to desert’.

² I cannot take refuge in this position: ‘I think I am under a moral necessity to regard all my duties *as if they were* commandments of God, but I’m not entitled to accept as a matter of theory that any such supreme being really exists.’ Feeling obliged to believe for purposes of •practice something that I see no ground for accepting as a •theoretical truth? I’m so far from doing this that I cannot even conceive the state of mind that those words seem to describe, except as a momentary half-wilful irrationality brought on by a spasm of philosophical despair.

couldn't be satisfactorily answered without a general examination of the criteria of true and false beliefs. Those who hold that the structure of physical science is really built out of conclusions logically inferred from self-evident premises may reasonably demand that any practical judgments claiming philosophical certainty should be based on an equally firm foundation. But if we find that in our supposed knowledge of

the world of nature we accept as true universal propositions that seem to be based on nothing but the facts that •we have a strong disposition to accept them and •they are indispensable to the systematic coherence of our beliefs, we'll find it harder to reject a similarly supported assumption in ethics, without opening the door to universal scepticism.